

STIC-Biotech/ChemLib

74293

From: Yu, Misook  
Sent: Wednesday, August 28, 2002 8:35 AM  
To: STIC-Biotech/ChemLib  
Subject: 09/900,147

Please search SEQ ID NO:1

Examiner Misook Yu, Ph.D.  
703-308-2454 (Phone)  
Art Unit 1642  
CM1-8E18 (Room)  
CM1-8E12 (Mail Box)

Edward Hart  
Technical Info. Specialist  
STIC/Biotech  
CMI 6B02 Tel: 305-9203

All search here

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 8/28/02  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: 1 \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: PR \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

Run on: August 28, 2002, 12:59:19 ; Search time 25.55 seconds  
(without alignments)  
250.521 Million cell updates/sec

Title: US-09-900-147-1  
Perfect score: 186  
Sequence: 1 KNIRRVYDALNVTLMNITSEKREIKIKGLPTNSA 37

Scoring table: BLOSUM62

Searched: 562222 aeqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

MAXIMUM DB seq length: 2000000000

**Post-processing: Minimum Match On**

Listing first 45 summaries

Database : SPTREMBL\_19: \*

```

1:  sp.archae.*
2:  sp.bacteria.*
3:  sp.fungi.*
4:  sp.human.*
5:  sp.invertebrate.*
6:  sp.mammal.*
7:  sp.mhc.*
8:  sp.organelle.*
9:  sp.phage.*
10: sp.plant.*
11: sp.rodent.*
12: sp.virus.*
13: sp.vertebrate.*
14: sp.unclassified.*
15: sp.virus.*
16: sp.bacterialp.*
17: sp.archaeip.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of raw result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No	Score	Query Match	Length	DB	ID	Description
1	186	100.0	290	11	Q9C47	Q9C47 nus musculi
2	186	100.0	396	11	Q9D27	Q9D27 nus musculi
3	186	100.0	416	4	Q9C28	Q9C28 homo sapien
4	189	90.9	381	5	Q9C27	Q9C27 brigia mala
5	186	89.2	363	5	Q4400	Q4400 drosophila
6	186	89.2	445	5	Q9V90	Q9V90 drosophila
7	184	88.2	233	4	Q9R54	Q9R54 homo sapien
8	164	61.2	385	10	Q9R52	Q9R52 arabisdopsis
9	151	81.2	413	10	Q9L27	Q9L27 arabisdopsis
10	136	73.1	288	10	Q9L25	Q9L25 arabisdopsis
11	136	73.1	292	10	Q9R73	Q9R73 arabisdopsis
12	133	71.5	261	10	Q9E71	Q9E71 trillium sp
13	87.5	44.4	287	5	Q9B18	Q9B18 canorhabdi
14	87.5	44.4	412	5	Q9B78	Q9B78 canorhabdi
15	72.5	39.0	281	6	Q9SK1	Q9SK1 macaca fasc
16	72	38.7	323	10	Q9S752	Q9S752 arabisdopsis

17	71.5	38.4	38.5	11	02CR4	09Cb4	mus musculus
16	69.5	37.4	43.9	13	02IB10	09B10	xenopus laevis
15	68.5	36.8	43.9	4	02B244	09B24	homo sapien
21	68.5	36.8	43.9	10	02SS52	09SS2	alcortiana
22	67.5	36.3	43.9	10	02PNS0	09PNS0	chenopodium
23	67.5	36.3	39.3	10	02FR8	09FR8	oryza sativa
22	67.5	36.3	43.6	10	02PR2	09PR2	oryza sativa
24	67.5	36.3	43.8	10	02ST5	09ST5	criticula
25	66.5	35.8	33.3	11	02PL0	09PL0	mus musculus
26	66.5	35.8	39.1	10	02PMS8	09PMS8	arabidopsis
27	66.5	35.8	46.0	10	02PVT1	09PVT1	arabidopsis
28	66.5	35.8	46.0	10	02PVT1	09PVT1	arabidopsis
29	66.5	35.8	46.0	10	02PVT1	09PVT1	arabidopsis
30	66.5	35.8	36.1	4	02G607	09G607	homo sapien
31	66.5	34.7	33.9	10	02SR10	09SR10	arabidopsis
32	64.5	34.7	34.2	5	02KX87	09KX87	caenorhabditis
33	64.5	34.7	43.1	10	02LE14	09LE14	daucus carota
34	63.5	34.1	39.6	10	02PV70	09PV70	arabidopsis
35	63.5	34.1	39.6	10	02ALX0	09ALX0	arabidopsis
36	63.5	34.1	42.6	10	02FC57	09FC57	arabidopsis
37	61.5	33.1	48.3	10	02P09	09P09	arabidopsis
38	61.5	33.1	48.3	10	02PMS8	09PMS8	arabidopsis
39	61.5	33.1	48.3	10	02PVT0	09PVT0	arabidopsis
40	61.5	33.1	51.6	10	02M454	09M454	arabidopsis
41	61.5	33.1	53.2	10	02S439	09S439	arabidopsis
42	59	31.7	65.7	5	02O619	09O619	caenorhabditis
43	59	31.7	46.3	16	02KVA4	09KVA4	vibrio cholerae
44	58	31.2	37.0	5	07O051	07O051	drosohila
45	54.5	29.3	77.1	5	01S380	01S380	caenorhabditis

## ALIGNMENTS

RESULT	1			
ID	09c127	PRELIMINARY	PRJ:	290 MA.
OC	09c127			
OC	01-JUN-2001 (TREMBLER, 17, Created)			
D7	01-JUN-2001 (TREMBLER, 17, Last sequence update)			
D7	01-DEC-2001 (TREMBLER, 19, Last annotation update)			
DE	TRANSCRIPTION FACTOR DP 1.			
GN	TFDP1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxId:100930;			
RP	RP			
NC	SEQUENCE FROM N.A.			
NC	SPRAIN=C518L/63; TISSUE=EMBRYO;			
RA	MIMLID=J1081660; PubMed=11217831;			
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Akawa T., Hara A., Funahashi Y., Kano H., Aichi J., Fukuda S.,			
RA	Atsuga K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamaneke I.,			
RA	Saito T., Okazaki Y., Oobori T., Bono H., Katsukawa T., Saito R.,			
RA	Rodota K., Matsuda H.A., Ashburner M., Barclay B., Casavant T.,			
RA	Pfeilschmann V., Gaestelund T., Gissi C., Ring B., Kochise H.,			
RA	Kuenli P., Lewis S., Matsuo Y., Mikiid I., Pesele O., Quecuenh J.,			
RA	Schriai L.M., Steadil P., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai R., Oitdo T., Furuo F., Aono H., Baldessari R., Barah G.,			
RA	Blake J., Boffelli D., Boujuna D., Carninci P., de Bonaldi M.F.,			
RA	Brownstein M.J., Bult C., Platner C., Fujita M., Gariboldi M.,			
RA	Guatinch S., Hill D., Hoffman J., Hume D.A., Kamaya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Nishimura J., Mazzarelli J., Momberts P.,			
RA	Mordone P., Ring B., Ringwald M., Rodriguez I., Saito M.,			
RA	Sasai H., Sato K., Schoenbach C., Seje T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Togo-Oka K., Wang K.H., Wette C., Wiltcher C., Winkler L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa T., Kaveji H., Kontauki S.,			
RA	Hayashizaki Y.,			
RT	Functional annotation of a full-length mouse cDNA collection.*;			
RL	Nature 409:685-690(2001).			
EMBL	AK013180; BAB26895.1; -			
OR	HSSP; Q14188; ICF7.			

DR MGD: MGI:101934; Tfdpl.  
 DR InterPro: IPR003316; E2F\_TDP.  
 DR Pfam: PF02319; E2F\_TDP; 1  
 DR SEQUENCE 390 AA; 32496 MW; AA3079A6FC4F1DD0 CRC64;

Query Match 100.0%; Score 186; DB 11; Length 290;  
 Best Local Similarity 100.0%; Pred. No. 3, 5e-18;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNIRRYDALNVTAMNITSEKKEIKVIGLPTNSA 37  
 DB 43 KNIRRYDALNVTAMNITSEKKEIKVIGLPTNSA 79

RESULT 2  
 ID 090297 PRELIMINARY; PRT; 396 AA.

AC 090297;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE TRANSCRIPTION FACTOR DP 1.

GN TFDPL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;

RA [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=THYMUS;  
 RA MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 Aizawa K., Izawa A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 Kadoya T., Ohsaki Y., Gotohori T., Bono H., Kasanuma T., Saito R.,  
 Pleischmann N., Cammerland T., Gissi C., King B., Kochiba H.,  
 Schmal L.M., Scandali F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 Sakai K., Ohtsuka T., Furuno M., Aono H., Balderelli R., Barsh G.,  
 Blake J., Bonfelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 Brownstein M.J., Bull C., Fletcher C., Fujita K., Gariboldi M.,  
 Guelincich S., Hill D., Hofmann M., Huie D.A., Kanaya M., Lee M.H.,  
 Lyons P., Marchionni L., Mashima J., Matzei J., Nomberts P.,  
 Nordone P., Ring B., Schoenbach C., Sessa F., Shibata Y., Storch K.-F.,  
 Sasaki H., Sato K., Schoenbach C., Sessa F., Shibata Y., Storch K.-F.,  
 Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,  
 Wyszewski A., Yoshida K., Hasegawa T., Kawa H., Kohno S.,  
 Henschel T.,  
 RA Functional annotation of a full-length mouse cDNA collection.\*;  
 RT Nature 409:685-690(2001).  
 DR EMBL: AK019997; BAB1395.1;  
 DR HSSP: Q14188; ICF7.  
 DR MGD: MGI:101934; Tfdpl.  
 DR InterPro: IPR003316; E2F\_TDP.  
 DR Pfam: PF02319; E2F\_TDP; 1.  
 DR SEQUENCE 396 AA; 43804 MW; 392AE2BD816522 CRC64;

Query Match 100.0%; Score 186; DB 11; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 4, 9e-18;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNIRRYDALNVTAMNITSEKKEIKVIGLPTNSA 37  
 DB 163 KNIRRYDALNVTAMNITSEKKEIKVIGLPTNSA 199

RESULT 3  
 ID 090298 PRELIMINARY; PRT; 416 AA.  
 AC 090298;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE HYPOTHEICAL 46.0 KDa PROTEIN (FRAGMENT).  
 GN DREFP434622.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 NCBI\_TaxID=9606;

Query Match 100.0%; Score 186; DB 4; Length 416;  
 Best Local Similarity 100.0%; Pred. No. 5, 1e-18;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNIRRYDALNVTAMNITSEKKEIKVIGLPTNSA 37  
 DB 148 KNIRRYDALNVTAMNITSEKKEIKVIGLPTNSA 184

RESULT 4  
 ID 090297 PRELIMINARY; PRT; 381 AA.

AC 090297;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE TRANSCRIPTION FACTOR DP1.  
 GN Brugia malayi.  
 OS Eukaryota; Metazoa; Nematoda; Chromadorea; Spiruridae; Filarioidea;  
 OC Onchocercidae; Brugia.  
 NCBI\_TaxID=6279;

RA [1]  
 RC SEQUENCE FROM N.A.  
 RP Brugia malayi OPI transcription factor.\*;  
 RT Submitted (JUL-2000) to the EMBL/GenBank/DOJ databases.  
 DR EMBL: AF284340; AAG12473.1;  
 DR HSSP: Q14188; ICF7.  
 DR InterPro: IPR003316; E2F\_TDP.  
 DR Pfam: PF02319; E2F\_TDP; 1.  
 DR SEQUENCE 381 AA; 43524 MW; 63FE58F7D4FD7655 CRC64;

Query Match 90.0%; Score 169; DB 5; Length 381;  
 Best Local Similarity 88.9%; Pred. No. 1, 2e-15;  
 Matches 32; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KNIRRYDALNVTAMNITSEKKEIKVIGLPTNS 36  
 DB 153 KNIRRYDALNVTAMNITSEKKEIKVIGLPTNS 188

RESULT 5  
 ID 044080 PRELIMINARY; PRT; 363 AA.  
 AC 044080;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE TRANSCRIPTION FACTOR DP (FRAGMENT).

GN DP OR C04654.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pharyngota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OR NCBI\_TaxID=7227;  
 (1)  
 RP SEQUENCE FROM N.A.  
 RA Koyman J., Whitaker A.J., Orr-Heaver T.L.,  
 RA Genes Dev. 0:0-0(1997).  
 DR EMBL: AF011367; AAB87765.1; --  
 DR HSSP: Q14188; 1CF7.  
 DR FLYBASE: F89N011765; DP.  
 DR InterPro: IPR003116; E2F\_TDP.  
 DR Pfam: PF02319; E2F\_TDP; 1.  
 FT MCL\_TER 1 1  
 FT SEQUENCE 363 AA; 4188 MW; 61DC66773BA43D42 CRC64;

Query Match 89.2% Score 166; DB 5; Length 363;  
 Best Local Similarity 86.1% Pred. No. 3e-15;  
 Matches 31: Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 KNIRRYDALVLMANNISKEKKEINWIGLPTNS 36  
 DB 111 KNIRRYDALVLMANNISKEKKEINWIGLPTNS 156

RESULT 6  
 OY6M0 PRELIMINARY: PRT: 445 AA.

AC OY6M0: 017472: 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE DP GENE PRODUCT (TRANSCRIPTION FACTOR E2F DP SUBUNIT).  
 GN DP OR C04654.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pharyngota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OR NCBI\_TaxID=7227;  
 (1)  
 RP SEQUENCE FROM N.A.  
 RA Adams W.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amaratunga C., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blake J.R., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Ball J.F., Abbey A., An H.-J., Andrews-Plankkuch C., Baldwin D.,  
 RA Bellow R.M., Bau A., Baxendale J., Bayraktaroglu L., Basalyk E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Botova D., Botchan N.R., Bouck J., Brodeur P., Broderick P.,  
 RA Butts K.C., Busan D.A., Butler H., Cadieu L.B., Davies P.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo J., Decher A., Deng I., Meyers A.D., Dev I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Dutilleul K.J., Evangelista C.C., Ferraz C., Fierstein S., Fleischmann W.,  
 RA Fowler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong P., Gottrell J.H., Gu T., Guan P., Harris M.,  
 RA Harris M.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbegan C.,  
 RA Jaitani M., Kaul S., Karp G.H., Ke I., Keshishian J.A., Ketchum K.A.,  
 RA Kimmel A.E., Kodira C.D., Kraft C., Kravitz S., Rupp D., Lai Z.,  
 RA Lako P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu K., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mironov G., Mitsuoka N.V., Mobarry C., Morris J., Noshell A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paczab J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puft V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shie B.C., Siden-Kiamos I., Slapton M., Stupski M.P., Satch T.,  
 RA Suter E., Spedding A.C., Stepien M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang L.-Y., Wasserman D.A., Weinlock G.M., Weissbach J.,  
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.B., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong P.W., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin C.M., Venter J.C.,  
 RA The genome sequence of Drosophila melanogaster.  
 RL Science 287:2185-2195(2000).  
 (2)  
 RP SEQUENCE OF 1-61 FROM N.A.  
 RA MEDLINE: 98078672; PubMed-9418962;  
 RA Duronio R.J., Bonnette P.C., O'Farrell P.H.,  
 RA Mutations of the Drosophila ddp, dezf, and cyclin E genes reveal  
 RA distinct roles for the E2F-DP transcription factor and cyclin E during  
 RA the G1-S transition.  
 RL Mol. Cell Biol. 18:141-151(1998).  
 DR EMBL: AF003819; MAF8403.1; --  
 DR EMBL: AF031700; MAF02971.1; --  
 DR HSSP: Q14188; 1CF7.  
 DR FLYBASE: F89N011765; DP.  
 DR InterPro: IPR003116; E2F\_TDP.  
 DR Pfam: PF02319; E2F\_TDP; 1.  
 FT SEQUENCE 445 AA; 49749 MW; 87F4DB826CFE4E42 CRC64;

Query Match 89.2% Score 166; DB 5; Length 445;  
 Best Local Similarity 86.1% Pred. No. 3e-15;  
 Matches 31: Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 KNIRRYDALVLMANNISKEKKEINWIGLPTNS 36  
 DB 213 KNIRRYDALVLMANNISKEKKEINWIGLPTNS 240

RESULT 7  
 OY6M24 PRELIMINARY: PRT: 233 AA.

AC OY6M24: 017472: 13, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE DP GENE PRODUCT (TRANSCRIPTION FACTOR E2F DP SUBUNIT).  
 GN DP OR C04654.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OR NCBI\_TaxID=9606;  
 (1)  
 RP SEQUENCE FROM N.A.  
 RA Pang X., Vaughan H., Chen M., Stclair F., Han K., Gebon J.,  
 RA "Cloning and characterization of novel genes related to hepatocellular  
 RA carcinoma." J. Hepatol. 28:1999-2000 (1999).  
 RL PubMed: 10819999; Medline: 10819999; PubMed: 10819999;  
 DR EMBL: AF219119; AAF73562.1; --  
 DR HSSP: Q14188; 1CF7.  
 DR InterPro: IPR003116; E2F\_TDP.  
 DR Pfam: PF02319; E2F\_TDP; 1.  
 FT SEQUENCE 233 AA; 26053 MW; C5F9F8D2B79E4784 CRC64;

Query Match 88.2% Score 164; DB 4; Length 233;  
 Best Local Similarity 86.5% Pred. No. 3e-15;  
 Matches 32: Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 KNIRRYDALVLMANNISKEKKEINWIGLPTNSA 37  
 DB 99 KNIRRYDALVLMANNISKEKKEINWIGLPTNSA 135

RESULT 8  
 O9LSE7 PRELIMINARY: PRT: 385 AA.  
 AC O9LSE7;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE DP-LIKE PROTEIN.  
 GN DPA.  
 OS Arabidopsis thaliana (mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-20562805; PubMed-1110847;  
 RA Mayar E., Atanasova A., de Veylder L., Rombauts S., Inze D.;  
 RT "Characterization of two distinct DP-related genes from Arabidopsis  
 thaliana."  
 RL FEBS Lett. 486:79-87(2000).  
 DR EMBL: AJ294532; CAC15484.1;  
 DR HSSP: Q14188; ICF7.  
 DR InterPro: IPR003316; E2F-TDP.  
 DR Pfam: PF02319; E2F-TDP; 1.  
 SO SEQUENCE 385 AA; 41755 MW; 5DB4ACAD0C52AF8 CRC64;

Query Match 81.2% Score 151; DB 10; Length 385;  
 Best Local Similarity 83.3%; Pred. No. 4, 1e-11;  
 Matches 30; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 KNIRRYDALNVLMNNISKEKIKVGLPTKS 36  
 DB 157 KNIRRYDALNVLMNNISKEKIKVGLPTKS 187

RESULT 9  
 O9LSE7 PRELIMINARY: PRT: 413 AA.  
 AC O9LSE7;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE TRANSCRIPTION FACTOR-LIKE PROTEIN.  
 GN P13B4\_160.  
 OS Arabidopsis thaliana (mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Pohl T., Weisenegger T., Bancroft I., Meyers H.W., Rudd S.,  
 RA Lencze K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RT EMBL: AL162751; CAB81399.1;  
 DR HSSP: Q14188; ICF7.  
 DR InterPro: IPR003316; E2F-TDP.  
 DR Pfam: PF02319; E2F-TDP; 1.  
 SO SEQUENCE 413 AA; 45604 MW; 239FDC6AED2F723 CRC64;

Query Match 81.2% Score 151; DB 10; Length 413;  
 Best Local Similarity 83.3%; Pred. No. 4, 5e-11;  
 Matches 30; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 KNIRRYDALNVLMNNISKEKIKVGLPTKS 36  
 DB 157 KNIRRYDALNVLMNNISKEKIKVGLPTKS 187

DB 169 KNIRRYDALNVLMNNISKEKIKVGLPTKS 204

RESULT 10  
 O9LSE5 PRELIMINARY: PRT: 288 AA.  
 AC O9LSE5;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE DP-2 TRANSCRIPTION FACTOR-LIKE.  
 GN T22P11\_60.  
 OS Arabidopsis thaliana (mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Quasthoff A.,  
 RA Bancroft I., Meyers H.W., Rudd S., Lencze K., Mayer K.F.X.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AL162971; CAB85984.1;  
 DR HSSP: Q14188; ICF7.  
 DR InterPro: IPR003316; E2F-TDP.  
 DR Pfam: PF02319; E2F-TDP; 1.  
 SO SEQUENCE 288 AA; 32560 MW; ECAD07DC06E92B7 CRC64;

Query Match 73.1% Score 136; DB 10; Length 288;  
 Best Local Similarity 75.8%; Pred. No. 4e-11;  
 Matches 25; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 KNIRRYDALNVLMNNISKEKIKVGLPTKS 33  
 DB 99 KNIRRYDALNVLMNNISKEKIKVGLPTKS 131

RESULT 11  
 O9LSE3 PRELIMINARY: PRT: 292 AA.  
 AC O9LSE3;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE DP-LIKE PROTEIN (EC 6.3.2.19).  
 GN DPA.  
 OS Arabidopsis thaliana (mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-20562805; PubMed-1110847;  
 RA Mayar E., Atanasova A., de Veylder L., Rombauts S., Inze D.;  
 RT "Characterization of two distinct DP-related genes from Arabidopsis  
 thaliana."  
 RL FEBS Lett. 486:79-87(2000).  
 DR EMBL: AJ294531; CAC15483.1;  
 DR HSSP: Q14188; ICF7.  
 DR InterPro: IPR003316; E2F-TDP.  
 DR Pfam: PF02319; E2F-TDP; 1.  
 SO SEQUENCE 292 AA; 33038 MW; 644324E13561FEC5 CRC64;

Query Match 73.1% Score 136; DB 10; Length 292;  
 Best Local Similarity 75.8%; Pred. No. 4, 1e-11;





Wed Aug 28 13:32:30 2002

us-09-900-147-1.rpx

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sv model

Run on: August 28, 2002, 12:57:49 : Search time 18.36 seconds  
(without alignments)  
193,644 Million cell updates/sec

Title: US-09-900-147-1

Sequence: 186  
1 KNIRRVYDALVLYAMNIISEKKEIKNIGLPTNSA 37

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: PIR\_71.0  
2: PIR\_1.0  
3: PIR\_2.0  
4: PIR\_3.0

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	186	100.0	410	2	transcription factor
2	186	100.0	410	2	transcription factor
3	186	100.0	410	2	transcription factor
4	186	100.0	410	2	transcription factor
5	186	100.0	410	2	transcription factor
6	186	100.0	410	2	transcription factor
7	186	100.0	410	2	transcription factor
8	186	100.0	410	2	transcription factor
9	186	100.0	410	2	transcription factor
10	186	100.0	410	2	transcription factor
11	186	100.0	410	2	transcription factor
12	186	100.0	410	2	transcription factor
13	186	100.0	410	2	transcription factor
14	186	100.0	410	2	transcription factor
15	186	100.0	410	2	transcription factor
16	186	100.0	410	2	transcription factor
17	186	100.0	410	2	transcription factor
18	186	100.0	410	2	transcription factor
19	186	100.0	410	2	transcription factor
20	186	100.0	410	2	transcription factor
21	186	100.0	410	2	transcription factor
22	186	100.0	410	2	transcription factor
23	186	100.0	410	2	transcription factor
24	186	100.0	410	2	transcription factor
25	186	100.0	410	2	transcription factor
26	186	100.0	410	2	transcription factor
27	186	100.0	410	2	transcription factor
28	186	100.0	410	2	transcription factor
29	186	100.0	410	2	transcription factor
30	186	100.0	410	2	transcription factor

30	53	28.5	334	2	A83991
31	53	28.5	610	2	T16194
32	52.5	28.2	292	2	G97111
33	52	28.0	369	2	G71510
34	52	28.0	388	2	I69644
35	52	28.0	967	2	G96537
36	51.5	27.7	901	2	D70116
37	51	27.4	307	2	G89953
38	51	27.4	395	2	D86876
39	51	27.4	1829	2	T34239
40	50.5	27.2	240	2	AH0224
41	50.5	27.2	282	2	T32145
42	50.5	27.2	419	2	E64526
43	50	26.9	385	2	G89956
44	50	26.9	456	2	D97304
45	50	26.9	600	2	T09676

#### ALIGNMENTS

##### RESULT 1

transcription factor DRTF1 chain 1 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Nov-1994 (sequence, revision 11-Nov-1994 text, change 10-Dec-1993)

C:Accession: B48585; S38372; S38371

R:Hein, K.; Wu, C.L.; Fetzney, A.R.; Lees, J.A.; Dynlacht, B.D.; Ngwu, C.; Harlow, Gene Dev. 7, 1850-1861, 1993

A:Title: Heterodimerization of the transcription factors E2F-1 and DP-1 leads to C

A:Reference number: A48585; MUID:94010284

A:Accession: B48585

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-410 <GIR>

A:Cross-references: EMBL:72310; MID:9395280; PION:CAAS1056.1; PID:9395281

R:Girling, R.; Partridge, J.E.; Bandara, L.R.; Burden, N.; Totty, M.F.; Newen, J.J

A:Title: A new component of the transcription factor DRTF1/E2F.

A:Reference number: S38371; MUID:94019777

A:Accession: S38371

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 188-410 <G12>

A:Cross-references: EMBL:72310

C:Superfamily: transcription factor DP

C:Keywords: DNA binding; transcription factor

Query Match 100.0% Score 186; DB 2; Length 410;

Best Local Similarity 100.0%; Pred. No. 3.3e-18;

Matches 37; Conservative 0; Mismatches 0; Indels 0;

1 KNIRRVYDALVLYAMNIISEKKEIKNIGLPTNSA 37  
|||||  
163 KNIRRVYDALVLYAMNIISEKKEIKNIGLPTNSA 199

##### RESULT 2

transcription factor DP-1 - human

C:Species: Homo sapiens (man)

C:Date: 11-Nov-1994 (sequence, revision 11-Nov-1994 text, change 10-Dec-1993)

C:Accession: A48585  
R:Hein, K.; Wu, C.L.; Fetzney, A.R.; Lees, J.A.; Dynlacht, B.D.; Ngwu, C.; Harlow, Gene Dev. 7, 1850-1861, 1993



A:Title: Heterodimerization of the transcription factors E2F-1 and DP-1 leads to cooperat

A:Accession: A48585

A:Status: Preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-410 <BL>

A:Cross-references: GB:U2959, NID:9414116, PIDM:AA58440.1, PID:9113117

C:Superfamily: transcription factor DP

C:Keywords: DNA binding; transcription factor

Query Match 100.0%; Score 186; DB 2; Length 410;

Best Local Similarity 100.0%; Pred. No. 5,3e-18;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 163 KNIRRVYDALVLMANNISKEKIKWGLPTNSA 159

RESULT 3

hypothetical protein DKF2P434G22.1 - human (fragment)

A:Alternate names: E2F dimerization partner 2

C:Species: Homo sapiens (man)

C:Date: 23-Jul-1999 sequence, revision 23-Jul-1999 text, change 17-Nov-2000

C:Accession: T12511; A57381; I37297; I39180

R:Amberg, M.; Mitter, U.; Meyer, H.W.; Gassenhuber, J.; Mennemann, S.

submitted to the Protein Sequence Database, June 1999

A:Reference number: 217527

A:Accession: T12511

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-416 <ANS>

A:Cross-references: EMBL:AL080206

A:Experimental source: adult testis; clone DKF2P434G22

R:Mu, C.L.; Zuberberg, L.R.; Ngwu, C.; Harlow, E.; Lees, J.A.

Mol. Cell. Biol. 15, 2536-2546, 1995

A:Title: In vivo association of E2F and DP family proteins.

A:Reference number: A57381; NID:9557935

A:Accession: A57381

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-416 <ANS>

A:Cross-references: GB:L40386; NID:9703084; PIDM:AA69016.1; PID:9703085

R:Zheng, Y.; Chelappan, S.P.

Oncogene 10, 2085-2093, 1995

A:Title: Cloning and characterization of human DP2, a novel dimerization partner of E2F

A:Reference number: I37297; NID:95303470

A:Accession: I37297

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-33-143 'O' 144-416 <ZNA>

A:Cross-references: EMBL:U18422; NID:9604478; PIDM:AA660378.1; PID:9604479

A:Accession: I39180

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-33-49 <ZNA>

A:Cross-references: EMBL:U51117; NID:91008545; PIDM:AMC50642.1; PID:91008546

C:Genetics:

A:Gene: GDB:TFDP2; DP2

A:Cross-references: CDB:638814

A:Map position: 1293-1293

A:Note: DKF2P434G22.1

C:Superfamily: transcription factor DP

Query Match 100.0%; Score 186; DB 2; Length 416;

Best Local Similarity 100.0%; Pred. No. 5,4e-18;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 KNIRRVYDALVLMANNISKEKIKWGLPTNSA 37

DB 148 KNIRRVYDALVLMANNISKEKIKWGLPTNSA 184

RESULT 4

transcription factor DTF1 chain 1 - mouse

A:Alternate names: transcription factor DP-1

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-1993 sequence, revision 10-Sep-1993 text, change 10-Dec-1999

C:Accession: S30049; S34572

R:Girling, R.; Partidge, J.F.; Bandaru, L.R.; Burden, N.; Totty, N.F.; Hansen, J.J.

Nature 362, 83-87, 1993

A:Title: A new component of the transcription factor DTF1/E2F.

A:Reference number: S30049; NID:93188967

A:Accession: S30049

A:Molecule type: mRNA

A:Residues: 1-429 <GIR>

A:Cross-references: EMBL:X72310

A:Accession: S34572

A:Molecule type: protein

A:Residues: 75-90; 131-131; 164-184; 191-207; 215-249; 261-293; 302-313; 321-328 <GIR>

C:Superfamily: transcription factor DP

C:Keywords: DNA binding; transcription factor

P184-204/Domain: DNA binding status predicted <DNA>

Query Match 100.0%; Score 186; DB 2; Length 429;

Best Local Similarity 100.0%; Pred. No. 5,6e-18;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 163 KNIRRVYDALVLMANNISKEKIKWGLPTNSA 199

DB 163 KNIRRVYDALVLMANNISKEKIKWGLPTNSA 199

RESULT 5

transcription factor DP - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 25-Aug-1995 sequence, revision 25-Aug-1995 text, change 10-Dec-1999

C:Accession: B55745

R:Dynlacht, B.D.; Brock, A.; Dembki, M.; Fennah, L.; Dyson, M.

Proc. Natl. Acad. Sci. U.S.A. 91, 6359-6363, 1994

A:Title: DNA-binding and trans-activation properties of Drosophila E2F and DP prote

A:Reference number: A55745; NID:94294181

A:Accession: B55745

A:Status: Preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-377 <DYN>

A:Cross-references: GB:X79708; NID:9516866; PID:9516867

C:Genetics:

A:Gene: flyBase:DP

A:Cross-references: flyBase:FBgn0011763

C:Superfamily: transcription factor DP

Query Match 89.3%; Score 166; DB 2; Length 377;

Best Local Similarity 86.1%; Pred. No. 2,9e-15;

Matches 31; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

DB 145 KNIRRVYDALVLMANNISKEKIKWGLPTNSA 180

RESULT 6

hypothetical protein T23207.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1998 sequence, revision 15-Oct-1999 text, change 15-Oct-1999

C:Accession: T23207

R:Barlow, K.

submitted to the EMBL Data Library, December 1995





Query Match 37.4%; Score 69.5; DB 2; Length 437;  
 Best Local Similarity 44.8%; Pred. No. 0.095;  
 Matches 13; Conservative 7; Mismatches 6; Indels 1; Gaps 1;  
 QY 4 RRRTYDALVYKANNISKE-KKEIKWIG 31  
 DB 166 KRRTIDITWLEGIOLIRKKAKNNIOWG 194

## RESULT 13

A56199  
 Description factor E2F-1 - fruit fly (Drosophila melanogaster)  
 Species: Drosophila melanogaster  
 Date: 28-Apr-1995; Revision 28-Apr-1995; Change 21-Jul-2000  
 Accession: A56199; A55745  
 Olcanil, K.; Nevins, J.R.  
 Mol. Cell. Biol. 14, 1603-1612, 1994  
 Title: Functional properties of a Drosophila homolog of the E2F1 gene.  
 Reference number: A56199; MID:94158832  
 Accession: A56199  
 Status: Preliminary  
 Molecule type: mRNA  
 Residues: 1-805 <ORF>  
 Cross-references: GB:U10184; MID:9499103; PIDN:AA19003.1; PID:9499104  
 Dynlacht, B.D.; Brook, A.; Dembki, M.; Yenush, L.; Dyson, N.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 6359-6363, 1994  
 Title: DNA-binding and trans-activation properties of Drosophila E2F and DP proteins.  
 Reference number: A55745; MID:94294381  
 Accession: A55745  
 Status: Preliminary; nucleic acid sequence not shown  
 Molecule type: mRNA  
 Residues: 1-805 <ORF>  
 Cross-references: GB:X78421; MID:9516868; PIDN:CAA55186.1; PID:9516869  
 Genetics:  
 Gene: FlyBase:E2F1  
 Cross-references: FlyBase:FBgn0011767

Query Match 36.3%; Score 67.5; DB 2; Length 805;  
 Best Local Similarity 48.1%; Pred. No. 0.36;  
 Matches 13; Conservative 6; Mismatches 7; Indels 1; Gaps 1;  
 QY 4 RRRTYDALVYKANNISKE-KKEIKW 29  
 DB 291 KRRTIDITWLEGIOLIRKKAKNNIOW 317

Search completed: August 28, 2002, 12:59:59  
 Job time: 130 sec

Wed Aug 28 13:32:29 2002

us-09-900-147-1.xag

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 12:55:14 ; Search time 29.9 Seconds

(without alignments)  
137.449 Million cell updates/sec

Title: US-09-900-147-1

Sequence: 1 ENIRRRYDALNVLNMTISKEKEIKWIGLPTNSA 37

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Listing first 45 summaries

Database:

1: /A/Geneseq\_013803:.\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:.\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:.\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:.\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:.\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:.\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:.\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:.\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:.\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:.\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:.\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:.\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:.\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:.\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:.\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:.\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:.\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:.\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:.\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:.\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:.\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:.\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	186	100.0	37	19	AAW30501
2	186	100.0	149	22	ABG05510
3	186	100.0	353	22	NA674072
4	186	100.0	410	15	AA854653
5	186	100.0	410	17	AA88212
6	186	100.0	424	22	ABG28057
7	182	97.8	446	20	AA178035
8	166	89.2	445	22	ABG08852
9	162	87.1	185	22	ABG03508
10	158	84.9	119	21	AA83108
11	153	83.3	120	21	AA831007

12	153	82.3	575	20	AAW81389
13	152	81.7	30	19	AAW30507
14	151	81.2	277	21	AA628852
15	151	81.2	299	21	AA607975
16	151	81.2	344	21	AA628851
17	151	81.2	385	21	AA628850
18	144	77.4	194	20	AA123164
19	135	73.1	138	20	AA123161
20	133	71.5	56	22	AA677365
21	133	71.5	261	22	AA677362
22	116	62.4	83	20	AA132163
23	101	54.3	20	19	AAW30505
24	101	54.3	28	19	AAW3051
25	101	54.3	28	19	AAW30515
26	91	48.9	19	19	AAW30504
27	83	44.6	19	19	AAW30515
28	80	43.0	15	19	AAW30510
29	79	42.5	19	19	AAW30517
30	77	41.4	19	19	AAW30516
31	76	40.9	16	19	AAW30506
32	74.5	40.1	272	18	AAW37447
33	72.5	39.0	281	18	AAW37489
34	72.5	39.0	281	20	AA108415
35	72.5	39.0	15	19	AAW37052
36	71.5	38.4	196	20	AAW37052
37	71.5	38.4	437	19	AAW37052
38	71.5	38.4	437	19	AAW37052
39	71.5	38.4	437	19	AAW37052
40	71.5	38.4	437	20	AAW37052
41	71.5	38.4	437	21	AAW37052
42	71.5	38.4	437	15	AAW37052
43	69.5	37.4	437	16	AAW37052
44	69	37.1	14	19	AAW37052
45	68.5	36.8	465	22	AAW37052

#### ALIGNMENTS

RESULT 1  
ID AAW30501 standard; Peptide: 37 AA.  
XX  
NC AAW30501:  
XX  
DT 26-OCT-1998 (first entry)  
XX  
DE DP-1 transcription factor peptide H (DEP box).  
XX  
XX DP-1 transcription factor; antagonist; E2F protein; apoptosis;  
XX cell proliferation; cardiovascular cell; restenosis; tumour;  
XX surgical stent; therapy.  
XX  
XX Synthetic.  
XX OS Homo sapiens.  
XX PM MO9828314-N1.  
XX  
XX 02-JUL-1998.  
XX  
XX 22-DEC-1997; 97MO-GB03506.  
XX  
XX 20-DEC-1996; 96GB-0026589.  
XX  
XX (PROL-) PROLIFIX LTD.  
XX  
XX Bandera LR, La Thangue NB;  
XX WPI: 1998-377596/32.  
XX  
XX Polypeptide fragments of the DP-1 transcription factor - used for  
XX inducing apoptosis, specifically in tumour and cardiovascular cells,  
XX e.g. for preventing restenosis

XX Claim 1: Page 44: 55pp: English.  
 XX  
 CC Peptide H (I) comprises residues 163-199, i.e. the DEF box region,  
 CC of transmembrane factor DFL. Claimed fragments (II) (see AA020502-07)  
 CC of (I) are capable of antagonizing the heterodimerization of a DP  
 CC protein with an EGF protein. Also claimed are fusion proteins  
 CC (III) comprising (I) or (II) and a membrane translocation sequence  
 CC (see AA03008), expression vectors encoding (I)-(III) and host cells  
 CC (I)-(III) are used therapeutically to induce apoptosis,  
 CC specifically in tumor or cardiovascular cells, either in vivo or in  
 CC vitro, e.g. for purging bone marrow. Surgical stents comprising  
 CC (I)-(III) are used to treat or prevent restenosis in patients who  
 CC have undergone angioplasty. (I)-(III) function by inactivating  
 CC the DNA-binding activity of DP/EGF heterodimers. They are also  
 CC used as research reagents, as positive controls in assays for  
 CC identifying antagonists of DP-1/EGF dimerization and as immunosassay  
 CC agents. Also described is the use of sequences antisense to  
 CC nucleic acids encoding (I)-(III) to control DP levels in cells,  
 CC or cytostatic agents, (I)-(III) enhance cell killing.  
 XX  
 XX Sequence 37 AA:

Query Match 100.0%; Score 186; DB 19; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 9, 8e-22;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KNTRRYDALVLTAMNITSEKREKNGIGPTNSA 37  
 DB 1 KNTRRYDALVLTAMNITSEKREKNGIGPTNSA 37

## RESULT 2

AB005510  
 ID AB005510 standard; Protein: 149 AA.

XX AB005510;

XX 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #5501.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN W0200175057-A2.

XX PD 11-OCT-2001.

XX PE 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540237.

XX PR 23-AUG-2000; 2000US-0619167.

XX PA (HUS-) HUSBD INC.

XX PI Dymac RT, Liu C, Tang YZ;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS69697.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
 XX diagnostics, forensics, gene mapping, identification of mutations  
 XX responsible for genetic disorders or other traits and to assess  
 XX biodiversity

XX PS Claim 20; SEQ ID NO 35869; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantifying a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AB00010-AB03037 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from HMO  
 CC at ftp://ftp.hmo.gov/pub/published\_pcr\_sequences.  
 XX  
 XX Sequence 149 AA:

Query Match 100.0%; Score 186; DB 22; Length 149;  
 Best Local Similarity 100.0%; Pred. No. 5, 3e-21;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KNTRRYDALVLTAMNITSEKREKNGIGPTNSA 37  
 DB 22 KNTRRYDALVLTAMNITSEKREKNGIGPTNSA 58

## RESULT 3

AG74072  
 ID AG74072 standard; Protein: 355 AA.

XX AG74072;

XX 03-SEP-2001 (first entry)

XX DE human colon cancer antigen protein SPO ID NO:4836.

XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
 XX colorectal carcinoma; chromosome 13.

XX OS Homo sapiens.

XX PN W0200122920-A2.

XX PD 05-APR-2001.

XX PE 28-SEP-2000; 2000WO-US26524.

XX PR 29-SEP-1999; 98US-035807.

XX PR 03-NOV-1999; 93US-0163780..

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;

XX DR WPI; 2001-23537/74.

XX DR N-PSDB; AAB33503.

XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
 XX useful for preventing, diagnosing and/or treating colorectal cancers -  
 XX Claim 11; Page 6620-6621; 9803pp; English.

XX AA032943 to AA037195 and AG773514 to AG77788 represent human colon  
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
 CC the proteins are collectively known as colon cancer antigens. The colon  
 CC cancer antigens have cytostatic activity and can be used in gene

CC therapy of vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders  
CC associated with decreased expression by rectifying mutations or deletions  
CC in a patient's genome that affect the activity of P by expressing  
CC inactive proteins or to supplement the patients own production of P.  
CC Additionally, N may be used to produce the colon cancer-associated Ps,  
CC by inserting the nucleic acids into a host cell and culturing the cell  
CC to express the proteins. N and P can be used in the prevention, diagnosis  
CC and treatment of colorectal carcinomas and cancers. AAM37704  
CC and AAM37789 represent sequences used in the exemplification of the  
CC present invention.

CC C.N.B., Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 355 Mb

```

Query Match 100.0% Score 165 DB 23 Length 355:
Best Local Similarity 100.0% Prd. No. 1, 5e-20:
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0

0Y 1 KRIIRRYDALNYLMAHHIISKKSKKINICLPPNS 37
      |||||
Db 169 KRIIRRYDALNYLMAHHIISKKSKKINICLPPNS 205

```

RESULT 4  
AAR54663  
ID AAR54663 standard; Protein: 410 AA  
xx

07 15-OCT-1994 (first entry)  
 06 Transcription factor DP-1.  
 05 Transcription factor: DP-1; E2F-1; proliferative disease;  
 04 cancer; gene therapy; P9-EC.  
 03  
 02 Kus sp.  
 01 M09410307-A.  
 00  
 59 11-MAY-1994.  
 58  
 57 29-OCT-1993; 93MO-GB02227.  
 56 29-OCT-1993; 92GB-0022715.  
 55 05-AUG-1993; 93GB-0016506.  
 54 (MED-) MEDICAL RES COUNCIL.  
 53  
 52 LA THANCUE NB:  
 51  
 50 MPI; 1994-167464/20.  
 49 M-PSDB; M065344.  
 48  
 47 Novel transcription factor DP-1 and its recombinant production  
 46 for diagnosis and treatment of diseases associated with abnormal  
 45 E2F or E2F-1 and family expression, esp. proliferative diseases  
 44  
 43 Disclosure: Page 49-50; 74pp: English.  
 42  
 41 DNA encoding novel transcription factor DP-1 was obtained from  
 40 P9-EC cells. The DNA can be used to produce recombinant  
 39 DP-1, or in gene therapy as a means of treating the uncontrolled  
 38 proliferation of e.g. tumor cells.  
 37  
 36  
 35  
 34  
 33  
 32  
 31  
 30  
 29  
 28  
 27  
 26  
 25  
 24  
 23  
 22  
 21  
 20  
 19  
 18  
 17  
 16  
 15  
 14  
 13  
 12  
 11  
 10  
 09  
 08  
 07  
 06  
 05  
 04  
 03  
 02  
 01  
 00  
 59  
 58  
 57  
 56  
 55  
 54  
 53  
 52  
 51  
 50  
 Sequence 410 AA:

```

Query Match      100.0%: Score 166, Db 15; length 410;
Best Local Similarity 100.0%, Pred. No. 1,8e-203
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

RESULT	5
AA089212	
ID	AA089212 standard; Protein; 410 AA

DT 16-JUL-1996 (first entry)  
XX  
DE Transcription factor DP-1.  
XX  
XX DP-1: transcription factor; DNA binding protein; cell cycle  
RV phosphorylation; cell proliferation; therapy; diagnosis.  
RW  
XX Homo sapiens.  
XS

	Location/Qualifiers
key	388
PH Modified-site	/label- Phosphorylation_site
PT Modified-site	/label- Phosphorylation_site
PI	
PK	
PN	
NO960125-12.	

PD 18-JAN-1996.

PF 03-JUL-1995; 95MO-GB01567

PR 01-JUL-1994; 94GB-0013327

PA (MEDI-) MEDICAL RES COUNCIL.

La Banque NB;

XX  
PR  
WPL: 1996-087752/09.

DK N-PSDB; AA710584.  
TX

PT Assays for potential

**PT** prevent or delay ent

XX

XX

complex. It is phos

CC Finding can be used

hypoparathyroidism. Such a

CC	cell cycle	info & p
CC	disorders	antibodies

phosphorylation (see

XX  
XX

**Query Match**

**Matches 37; Conserve**

by 1 KNIRRVYDALNVIL

Db 163 koltyvdalovj

ID	AAV18026	standard; Protein; 446 AA.
AC	AAV18026;	
XX	02-AUG-1999	(first entry)
DE	Murine DP-3alpha isoform protein sequence.	
KW	DIP: DP-interacting protein; B7B/POZ domain; transcriptional repressor; DP-dependent transcription; immunogen; cycle control; seminoma; teratoma; diagnosis; meiosis promoter; proliferative disorder; diagnosis.	
OS	Mus sp.	
PN	MO992709L.A1.	
PD	03-JUN-1999.	
PF	23-NOV-1998; 98MO-GDB3485.	
PR	21-NOV-1997; 97GB-0024828.	
PX	(UNIT) UNIV GLASGOW.	
PI	De La Lanza S, La Thangue NB;	
DR	WPI: 1999-147716/29.	
DT	N-PSTDB; AAV6730.	
PT	New DP transcription factor-interacting protein designated DIP	
PS	Disclosure; Page 73-76; 79pp; English.	
CC	This sequence is the murine DP-3alpha isoform. The invention relates to the DP-interacting protein (DIP), that contains a B7B/POZ domain, and is a potent transcriptional repressor that can inactivate DP-dependent transcription. The DIP polypeptides may be used as an immunogen or to raise antibodies. Such antibodies are useful in purification and diagnostic screening methods. DIP polypeptides may be used to screen for molecules which modulate its activity. Such molecules may be useful in therapeutic contexts, e.g. in connection with conditions which involve abnormal or aberrant expression of DIP. Modulators of DIP may be used to promote cell cycle control or modification in conditions such as seminomas and teratomas, and in the promotion of meiosis. DIP polynucleotides are useful as sources of probes and primers, and for screening for the presence of alleles, mutants, variants and polymorphisms. DIP polynucleotides may also be useful for the diagnosis of conditions associated with disorders of germ cells, including proliferative disorders such as seminomas and teratomas.	
XX	Sequence 446 AA;	
SQ	446	
Query Match	97.8%; Score 182; DB 20;	Length 446;
Best Local Similarity	97.3%; Pred. no. 8.4e-20;	
Matches	36; Conservative 1; Mismatches 0;	Gaps 0;
OY	1 KNIRBYDALNVAMNIIISREKRVICGLPTNS 37	
Db	178 enrrrygdalnvmnnllskkkelwvlpnns 216	
RESULT	8	
AB60852	AB60852 standard; Protein; 445 AA.	
XX	AB60852;	
XX	26-MAR-2002 (first entry)	



```

Query Match          89.2%  Score 166;  DB 22;  Length 445;
Best Local Similarity 86.1%  Pred. No. 2,6e-17;
Matches 31;  Conservative 4;  Mismatches 1;  Indels 0;  Gaps 0;

Oy      1  KNIRRYDALVLTAMNIIISREKKEINIGLPTNS 36
|||||
Db      213 KNIRRYDALVLTAMNIIISREKKEINIGLPTNS 248

RESULT  9
ABG05508
10  ABG05508 standard; Protein; 185 AA.
XK
XK  AC
XK  ABG05508;
D7      13-FEB-2002 (first entry)
XK
XK  D8
XK  Novel human diagnostic protein p5499.
XK
XK  Human; chromosome mapping; gene mapping; forensic;
XK  food supplement; medical imaging; diagnostic; genetic disorder.
O5      Homo sapiens.
XK
XK  PN
XK  WC0300175067-A2.
XK
XK  PD
XK  11-OCT-2001.
XK
XK  30-MAR-2001; 2001WO-US08631.
XK
XK  P

```

Query Match	Score	DB	Length
Best Local Similarity	89.28	Pred. No. 3,8e-17	
Matches	33	Conservative	2
		Mismatches	2
		Indels	0
		Gaps	0
1 KNIRRYDALNYLANNINISKEKEIKVIGLPTMSA 37			
:   :			
Db 149 knlrrrcdclnrlrrvslskkktkvlglptmsa 185			
RESULT 10			
AAB33108			
ID AAB33108 standard; Protein: 119 AA.			
XX			
AC AAB33108;			
XX			
DT 25-JAN-2001 (first entry)			
XX			
DE Pinus radiata transcription factor protein sequence #235.			
XX			
RM Plant: transcription factor; gene expression; eucalyptus; pine; acetabular; sweetgum; oak; mahogany; b1p; g-box binding factor; basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; homeodomain zipper; LIM domain; AP2; GRB3; zinc finger domain; type 2 cy2h1a2; CCAAT box element; MYB.			
XX			
OS Pinus radiata.			
XX			
PN MO200053724-A2.			
XX			
14-SEP-2000.			
XX			

XX 09-MAR-2000: 2000MO-US06112.  
 PF 11-MAR-1999: 99US-0266513.  
 XX 18-AUG-1999: 99US-0149485.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
 XX Wood M, McGrath A, Shank M, Glenn M  
 PI WPI: 2000-579369/54.  
 DR  
 XX New isolated polynucleotide encoding a plant transcription factor for  
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 PT having modified gene expression or modified activity of a polypeptide  
 PT  
 XX  
 PS Claim 8: Pages 430-431: 747pp: English.  
 CC The present invention relates to novel plant transcription factors from  
 CC Eucalyptus grandis or Pinus radiata. The present sequence is one such  
 CC transcription factor. The transcription factor may be used to produce a  
 CC plant having modified gene expression such as a woody plant e.g. a  
 CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or  
 CC to modify the activity of a polypeptide in a plant. The transcription  
 CC factors of the present invention are members from the following families  
 CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic  
 CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain  
 CC zipper, LIM domain, AP2 and ERBBs, zinc finger domains of type 2  
 CC CysHis2, CCNAF box elements and MYB.  
 CC  
 XX Sequence 119 AA:  
 SO  
 Query Match 84.98: Score 158: DB 21: Length 119:  
 Best Local Similarity 86.18: Pred. No. 9.4e-17:  
 Matches 31: Conservative 3: Mismatches 2: Indels 0: Gaps 0:  
 07 1 KNIRRYDALNVLNMMNISKKEIKWIGLPTNS 36  
 DB 71 knlrryvalnvlmmnlslskkkelrvgipld 106  
 RESULT 11  
 ID AAB33007 standard; Protein: 120 AA.  
 AC AAB33007:  
 XX  
 DT 25-JAN-2001 (first entry)  
 DE Pinus radiata transcription factor protein sequence 1134.  
 XX  
 XX Plant: transcription factor; gene expression; eucalyptus; pine; acacia;  
 KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
 KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
 KW homeodomain zipper; LIM domain; AP2; ERBBs; zinc finger domain;  
 KW type 2 CysHis2; CCNAF box element; MYB.  
 XX  
 OS Pinus radiata.  
 XX  
 PN MO200053724-A2.  
 PD 14-SEP-2000.  
 XX  
 XX 09-MAR-2000: 2000MO-US06112.  
 XX 11-MAR-1999: 99US-0266513.  
 XX 18-AUG-1999: 99US-0149485.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX Wood M, McGrath A, Shank M, Glenn M;  
 PI WPI: 2000-579369/54.  
 DR  
 XX New isolated polynucleotide encoding a plant transcription factor for  
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 PT having modified gene expression or modified activity of a polypeptide  
 PT  
 XX  
 PS Claim 8: Page 390: 747pp: English.  
 CC The present invention relates to novel plant transcription factors from  
 CC Eucalyptus grandis or Pinus radiata. The present sequence is one such  
 CC transcription factor. The transcription factor may be used to produce a  
 CC plant having modified gene expression such as a woody plant e.g. a  
 CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or  
 CC to modify the activity of a polypeptide in a plant. The transcription  
 CC factors of the present invention are members from the following families  
 CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic  
 CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain  
 CC zipper, LIM domain, AP2 and ERBBs, zinc finger domains of type 2  
 CC CysHis2, CCNAF box elements and MYB.  
 CC  
 XX Sequence 120 AA:  
 SO  
 Query Match 87.38: Score 153: DB 21: Length 120:  
 Best Local Similarity 83.38: Pred. No. 5.7e-16:  
 Matches 30: Conservative 4: Mismatches 2: Indels 0: Gaps 0:  
 07 1 KNIRRYDALNVLNMMNISKKEIKWIGLPTNS 36  
 DB 70 knlrryvalnvlmmnlslskkkelrvgipld 105  
 RESULT 12  
 ID AAB33089 standard; Protein: 575 AA.  
 AC AAB33089:  
 XX  
 DT 25-MAR-1999 (first entry)  
 DE Caenorhabditis elegans synxv protein LIN-55.  
 XX  
 XX LIN-55: synthetic multivulva; Synxv; signal transduction;  
 KW animal model; tumour suppressor; retinoblastoma; cancer; cancer;  
 KW cell proliferation; gene therapy.  
 XX  
 OS Caenorhabditis elegans.  
 XX  
 PN MO9854299-A1.  
 PD 03-DEC-1996.  
 XX  
 XX 26-MAY-1996: 98MO-US11043.  
 XX 26-MAY-1997: 97US-0047996.  
 XX (MASI) MASSACHUSETTS INST TECHNOLOGY,  
 PA Ceol C, Horvitz HR, Lu X;  
 PI WPI: 1999-045362/04.  
 DR N-PSDB: AAV72862.  
 XX  
 XX Novel LIN-37, -35, -55, -52, -53 and -54 gene from C. elegans -  
 PT useful for treating diseases associated with altered levels of cell  
 PT proliferation, e.g. carcinomas  
 XX  
 XX Claim 7: Fig 8: 70pp: English.





PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160880.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 23-OCT-1999; 99US-0161404.  
 PR 23-OCT-1999; 99US-0161405.  
 PR 23-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 81.2%; Score 151; DB 21; Length 277;  
 Best Local Similarity 83.3%; Pred. No. 3.2e-15;  
 Matches 30; Conservative 3; Mismatches 3; Indels 0; Caps 0;  
 Oy 1 KNIRRYVDALNVLAMNIISEKKEIRMGIPGTS 35  
 DB 44 KNIRRYVDALNVLAMNIISEKKEIRMGIPGTS 79

RESULT 15  
 AAB07975  
 ID AAB07975 standard; Protein: 299 AA.  
 AC AAB07975;  
 DT 14-NOV-2000 (first entry)  
 DE A corn transcription factor designated DP.  
 KW DP polypeptide; transcription factor; gene transcription; cell cycle;  
 KW DNA replication; DRR; differentiation-regulated transcription factor 1;  
 KW transgenic plant; transformation efficiency.  
 OS Zea mays.  
 TX  
 TX Key Location/Qualifiers  
 TX Misc-difference 1 /note="any amino acid encoded by NNN"  
 TX MO200047614-AA.  
 TX  
 TX 17-AUG-2000.  
 TX  
 TX 11-FEB-2000; 2000MO-US03651.  
 TX  
 TX 12-FEB-1999; 99US-0119857.  
 TX  
 TX (PION-) PIONEER HI-BRED INT INC.  
 TX  
 TX Lowe KS, Gordon-kamm NJ, Bailey MA, Gregory CA;  
 TX  
 TX N-PSDB; AAS59702.  
 TX  
 TX Transgenic plants comprising an expression cassette consisting of a DP  
 TX nuclear acid, when expressed, the nuclear acid increases the efficiency  
 TX of plant cell transformation by increasing the number of dividing cells  
 TX in the plant.  
 TX  
 TX Disclosure: Page 71-73; 76pp; English.

XX The present sequence represents a corn DP (not defined) polypeptide.  
 CC The polypeptide is a transcription factor that activates transcription  
 CC of numerous genes involved in DNA replication, thus playing a role in the  
 CC G1 to S transition in the cell cycle. Do, also called DRR  
 CC (differentiation-regulated transcription factor 1) has been shown to  
 CC form specific multiprotein complexes with the retinoblastoma  
 CC susceptibility protein, p107, cyclins and cdk2. The DP polynucleotide  
 CC was used to produce a transgenic plant. The DP nucleic acids and  
 CC proteins are useful for increasing transformation efficiency in plants  
 CC by increasing the number of dividing cells in the plant (dividing  
 CC cells may be more receptive to transformation).  
 CC  
 CC Sequence 299 AA;  
 SQ

Query Match 81.2%; Score 151; DB 21; Length 299;  
 Best Local Similarity 83.3%; Pred. No. 3.5e-15;  
 Matches 30; Conservative 3; Mismatches 3; Indels 0; Caps 0;  
 Oy 1 KNIRRYVDALNVLAMNIISEKKEIRMGIPGTS 36  
 DB 98 KNIRRYVDALNVLAMNIISEKKEIRMGIPGTS 133

Search completed: August 28, 2002, 12:59:14  
 Job time: 240 sec

Wed Aug 28 13:32:29 2002

us-09-900-147-1.rag

Page 10

Wed Aug 28 13:32:30 2002

us-09-900-147-1.rsp

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

On protein - protein search, using sw model

Run on: August 28, 2002, 12:59:34 ; Search time 10.33 Seconds

(without alignments)  
138,686 Million cell updates/sec

Title: US-09-900-147-1

Sequence: 186  
1 RNRIRRYDALMTVMAMNITSKREKRIWGLPTNSA 37

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 3819550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database: SWISSPROT\_40.0  
Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	186	100.0	410	1	TOP1_HUMAN
2	186	100.0	410	1	TOP1_HUMAN
3	186	100.0	446	1	TOP2_HUMAN
4	166	89.2	377	1	TOP2_MOUSE
5	152	82.3	598	1	TOP2_MOUSE
6	98.5	53.0	446	1	TOP2_MOUSE
7	74.5	40.1	272	1	TOP2_MOUSE
8	72.5	39.0	281	1	TOP2_MOUSE
9	71.5	38.4	430	1	TOP2_MOUSE
10	70.5	37.9	437	1	TOP2_MOUSE
11	69.5	37.4	437	1	TOP2_MOUSE
12	68.5	36.8	356	1	TOP2_MOUSE
13	67.5	36.3	405	1	TOP2_MOUSE
14	66.5	35.8	500	1	TOP2_MOUSE
15	65.5	35.3	335	1	TOP2_MOUSE
16	64.5	34.8	346	1	TOP2_MOUSE
17	63.5	34.3	413	1	TOP2_MOUSE
18	62.5	33.8	206	1	TOP2_MOUSE
19	61.5	33.3	369	1	TOP2_MOUSE
20	60.5	32.8	369	1	TOP2_MOUSE
21	59.5	32.3	369	1	TOP2_MOUSE
22	58.5	31.8	369	1	TOP2_MOUSE
23	57.5	31.3	369	1	TOP2_MOUSE
24	56.5	30.8	369	1	TOP2_MOUSE
25	55.5	30.3	369	1	TOP2_MOUSE
26	54.5	29.8	369	1	TOP2_MOUSE
27	53.5	29.3	369	1	TOP2_MOUSE
28	52.5	28.8	369	1	TOP2_MOUSE
29	51.5	28.3	369	1	TOP2_MOUSE
30	50.5	27.8	369	1	TOP2_MOUSE
31	49.5	27.3	369	1	TOP2_MOUSE
32	48.5	26.8	369	1	TOP2_MOUSE
33	47.5	26.3	369	1	TOP2_MOUSE
34	46.5	25.8	369	1	TOP2_MOUSE
35	45.5	25.3	369	1	TOP2_MOUSE
36	44.5	24.8	369	1	TOP2_MOUSE
37	43.5	24.3	369	1	TOP2_MOUSE
38	42.5	23.8	369	1	TOP2_MOUSE
39	41.5	23.3	369	1	TOP2_MOUSE
40	40.5	22.8	369	1	TOP2_MOUSE
41	39.5	22.3	369	1	TOP2_MOUSE
42	38.5	21.8	369	1	TOP2_MOUSE
43	37.5	21.3	369	1	TOP2_MOUSE
44	36.5	20.8	369	1	TOP2_MOUSE
45	35.5	20.3	369	1	TOP2_MOUSE

#### ALIGNMENTS

RESULT	TOP1_HUMAN	STANDARD	PRT	410 AA.
AC	014186			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	15-OCT-2001 (Rel. 40, Last annotation update)			
DE	Transcription factor DP-1 (E2F dimerization partner 1) (DNTP1-1)			
DE	polypeptide-1 (DNTP1-1)			
GN	DP-1 OR DP-1			
OS	Human sapiens (human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
OX	NCBI_TaxID=9606;			
RA	SEQUENCE FROM M.A. AND CHARACTERIZATION.			
RA	MEDLINE=94010284; PUBMED=8405993.			
RA	Hellin K., Mu C.-L., Fattley A.R., Lees J.A., Dynlacht B.D.,			
RA	"Heterodimerization of the transcription factors E2F-1 and DP-1 leads			
RA	to cooperative transcription."			
RA	Genes Dev. 7:1650-1661(1993).			
RA	[2]			
RA	CHARACTERIZATION.			
RA	MEDLINE=9525793; PUBMED=7739537.			
RA	Mu C.-L., Zukerberg L.R., Nyeu C., Harlow E., Lees J.A.,			
RA	"In vivo association of E2F and DP family proteins."			
RA	Mol. Cell. Biol. 15:2536-2546(1995).			
RA	[3]			
RA	PHOSPHORYLATION.			
RA	MEDLINE=94313991; PUBMED=8039504.			
RA	Bandaru L.R., Lam E.H.-T., Soehrensen T.S., Lammiman M., Gilling R.,			
RA	la Thangue N.B.,			
RA	"DP-1: a cell cycle-regulated and phosphorylated component of			
RA	transcription factor DNTP1/22 which is functionally important for			
RA	recognition by PRD and the adenovirus E4 orf 6/7 protein."			
RA	EMBO J. 13:3104-3114(1994).			
RA	[1]			
RA	FUNCTIONALITY: CAN STIMULATE E2F-DEPENDENT TRANSCRIPTION. BINDS DNA			
RA	COOPERATIVELY WITH E2F FAMILY MEMBERS THROUGH THE E2 RECOGNITION			
RA	SITE, TTTCC/GGCG, FOUND IN THE PROMOTER REGION OF A NUMBER OF			
RA	GENES WHOSE PRODUCTS ARE INVOLVED IN CELL CYCLE REGULATION OR IN			
RA	DNA REPLICATION. THE DP1/E2F COMPLEX FUNCTIONS IN THE CONTROL OF			
RA	CELL-CYCLE PROGRESSION FROM G1 TO S PHASE. THE E2F-1/DP COMPLEX			
RA	APPEARS TO MEDIATE BOTH CELL PROLIFERATION AND APOPTOSIS.			
RA	[1]			
RA	SUBUNIT: COMPONENT OF THE E2F/DP TRANSCRIPTION FACTOR COMPLEX.			
RA	BINDS HETEROLOGOUSLY WITH E2F FAMILY MEMBERS. THE COMPLEX CAN			
RA	INTERACT WITH HIPPOPHOSPHORYLATED RETINOBLASTOMA PROTEIN RB1 AND			
RA	RELATED PROTEINS (P107 AND P130) THAT INHIBIT THE E2F			
RA	TRANSACTIVATION DOMAIN. THIS REPRESSION INVOLVES RECRUITMENT OF			
RA	HISTONE DEACETYLASE (HDAC). DURING THE CELL CYCLE, FROM MID-TO-			
RA	LATE G1 PHASE, RB FAMILY MEMBERS BECOME PHOSPHORYLATED, DETACH			
RA	FROM THE DNTP1/E2F COMPLEX TO REVEAL E2F TRANSCRIPTIONALLY ACTIVE.			
RA	VIRAL ONCOPROTEINS, NOTABLY E1A, T-ANTIGEN AND HPV E7, ARE CAPABLE			
RA	OF SEQUESTERING RB PROTEIN, THUS RELEASING THE ACTIVE COMPLEX.			
RA	[1]			
RA	SUBCELLULAR LOCATION: Nucleus.			
RA	[1]			
RA	TISSUE SPECIFICITY: HIGHEST LEVELS IN MUSCLE. ALSO EXPRESSED IN			

NA DeAngelo L., Fry C.J., Bartley S.W., Farnham P., Gallie B.L.,  
RA Phillips R.A.:  
RT Expression patterns of the E2F family of transcription factors during  
RT mouse nervous system development. *J*  
RT Mech. Dev. 66:13-25(1997).  
RL (4)  
RP DEVELOPMENTAL EXPRESSION:  
RA DeAngelo L., Fry C.J., Bartley S.W., Farnham P., Gallie B.L.,  
RA Phillips R.A.:  
RT Expression patterns of the E2F family of transcription factors  
RT during murine epithelial development. *J*  
RL Cell Growth Differ. 8:553-563(1997).  
CC -1- FUNCTION: CAN STIMULATE E2F-DEPENDENT TRANSCRIPTION. BINDS DNA  
CC COOPERATIVELY WITH E2F FAMILY MEMBERS THROUGH THE E2 RECOGNITION  
CC SITE, TTTC/GCGC, FOUND IN THE PROMOTER REGION OF A NUMBER OF  
CC GENES WHOSE PRODUCTS ARE INVOLVED IN CELL CYCLE REGULATION OR IN  
CC DNA REPLICATION. THE b2F/E2F COMPLEX FUNCTIONS IN THE CONTROL OF  
CC CELL CYCLE PROGRESSION FROM G1 TO S PHASE. THE E2F-1/DP COMPLEX  
CC APPEARS TO MEDiate BOTH CELL PROLIFERATION AND APOPTOSIS.  
CC -1- SUBUNIT: COMPONENT OF THE DRP1/E2F TRANSCRIPTION FACTOR COMPLEX.  
CC FORMS HETERODIMERS WITH E2F FAMILY MEMBERS. THE COMPLEX CAN  
CC INTERACT WITH HYPOPHOSPHORYLATED RETINOBLASTOMA PROTEIN RB1 AND  
CC RELATED PROTEINS (P107 AND P130) THAT INHIBIT THE E2F  
CC TRANSCRIPTIONAL DOMAIN. DURING THE CELL CYCLE, RB BECOMES  
CC PHOSPHORYLATED IN MID-TO-LATE G1 PHASE, DETACHES FROM THE  
CC DRP1/E2F COMPLEX RENDERING E2F TRANSCRIPTIONALLY ACTIVE. VITAL  
CC ONCOPROTEIN, NOTABLY E1A, T-ANTIGEN AND HPV E7, ARE CAPABLE OF  
CC SEQUESTERING RB PROTEIN, THUS RELEASING THE ACTIVE COMPLEX.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- DEVELOPMENTAL STAGE: IN THE INTESTINAL EPITHELIUM, FIRST EXPRESSED  
CC IN UNDIFFERENTIATED AND MESCHERIAL TISSUES, LEVELS INCREASING BY  
CC 12.5 DPC IN THE EPITHELIAL COMPARTMENT WITH EPITHELIAL  
CC DIFFERENTIATION AT 15.5 DPC. DP-1 EXPRESSION INCREASES  
CC SUBSTANTIALLY IN THE INTRAVILLUS EPITHELIUM WITH LOWER LEVELS IN  
CC THE MESSCHERIAL. AT LATER STAGES, EXPRESSION COMBINES IN THE  
CC INTERVALS EPITHELIUM. ALSO FOUND AT LOWER LEVELS IN THE  
CC DEVELOPING VILI. IN THE DEVELOPING BRAIN, HIGHEST LEVELS FOUND  
CC BETWEEN 11.5 AND 13.5 DPC IN THE VENTRICULAR REGION. IN THE  
CC DEVELOPING RETINA, DP-1 IS EXPRESSED BOTH IN RETINOBLAST AND  
CC GANGLION CELL LAYERS FROM 14.3 DPC TO 6 DAYS AFTER BIRTH. IN OTHER  
CC DEVELOPING TISSUES, DP-1 EXPRESSION IS HIGHEST IN THE THYMUS. ALSO  
CC PRESENT IN KIDNEY, LUNG, LIVER, HEART AND CHONDROCYTES. WEAKLY  
CC EXPRESSED IN SKELETAL MUSCLE AND ABSENT FROM CHOROID PLEXUS.  
CC -1- PTM: PHOSPHORYLATION BY E2F-1-BOUND CYCLIN A-CDK2, IN THE S PHASE,  
CC INHIBITS E2F-MEDIATED DNA BINDING AND TRANSCRIPTION.  
CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.  
CC .....  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-ebc.ch/announce/>  
CC or send an email to [license@isb-ebc.ch](mailto:license@isb-ebc.ch)).  
CC .....  
DR EMBL: X7310; CAA51056.1; ..  
DR HSPB: 014188; ICF7.  
DR MGD: MGI:101934; Tfcdp1.  
DR InterPro: IPR003316; E2F\_TDP.  
DR Pfam: PF00319; E2F\_TDP\_1.  
KM Transcription regulation; DNA-binding; Activator; Nuclear protein;  
KM Phosphorylation; Cell cycle.  
FT DNP\_BIND 111 135 POTENTIAL.  
FT DNP\_BIND 161 135  
FT DNP\_BIND 204 277  
FT DNP\_BIND 214 246 DIMERIZATION (POTENTIAL).  
FT DNP\_BIND 259 315 DCB1.  
FT DNP\_BIND 250 253 DCB2.  
FT DNP\_BIND 395 410 POLY-PRO.  
FT DNP\_BIND 410 410 ASP/GLU-RICH (ACIDIC) (NCS DOMAIN).  
FT DNP\_BIND 23 23 PHOSPHORYLATION (BY CDK2) (POTENTIAL).  
SQ MOO\_RES 23 23  
SQ SEQUENCE 410 AA; 45330 MW; FFA090C76107189 CRC64;



Query Match 100.0%; Score 186; DB 1; Length 410;  
 Best Local Similarity 100.0%; Pred. No. 9, 2e-15;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KNIRRYVDALVLMANNISKEKKEIKWGLPTNSA 37  
 |||||||||||||||||||||||||||||||||||  
 DB 163 KNIRRYVDALVLMANNISKEKKEIKWGLPTNSA 169

RESULT 3  
 TDP2\_HUMAN STANDARD; PRT; 446 AA.  
 AC Q14188; Q14187; Q13331;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Transcription factor DP-2 (E2F dimerization partner 2).  
 GN TDP2 OR DP2  
 OS Homo sapiens (human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 NC NB1\_Tasid-9606;  
 RX SOURCE FROM N.A. (DELTA ISOFORM), AND CHARACTERIZATION.  
 RX MEDLINE-95357935; PubMed-779537;  
 RX Hu C.L., Sukerberg L.R., Ngwu C., Harlow E., Lees J.A.;  
 RX "In vivo association of E2F and DP family proteins.";  
 RX Mol. Cell. Biol. 15:2536-2546(1995).  
 RL [2]  
 RP SEQUENCE FROM N.A. (DELTA ISOFORM).  
 RP TISSUE-Kidney; PubMed-7784053;  
 RP MEDLINE-9530370; PubMed-7784053;  
 RP Zhang Y., Chellappan S.P.;  
 RP Cloning and characterization of human DP2, a novel dimerization  
 RP partner of E2F.";  
 RP Oncogene 10:7085-7093(1995).  
 RL [3]  
 RP SEQUENCE OF 1-102 FROM N.A. (ALPHA ISOFORM).  
 RP Zhang Y., Chellappan S.P.;  
 RP "Transcriptional activation and expression of DP transcription  
 RP factor during cell cycle and TPA-induced U937 differentiation.";  
 RP Submitted (DEC-1996) to the EMBL/GenBank/DBS databases.  
 RL [4]  
 RP SEQUENCE OF 63-79 FROM N.A. (DELTA ISOFORM), AND ALTERNATIVE SPLICING.  
 RP TISSUE-NerveTissue;  
 RP MEDLINE-96353862; PubMed-8755520;  
 RP Rogers K.T., Higgins P.D.R., Milla M.M., Phillips R.S., Horowitz J.M.;  
 RP "DP-2, a heterodimeric partner of E2F: identification and  
 RP characterization of DP-2 proteins expressed in vivo.";  
 RP Proc. Natl. Acad. Sci. U.S.A. 93:7594-7599(1996).  
 RL [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 121-215.  
 RP MEDLINE-99190701; PubMed-10090723;  
 RP Zheng M., Freeman E., Pabo C.O., Pavletich N.P.;  
 RP "Structural basis of DNA recognition by the heterodimeric cell cycle  
 RP transcription factor E2F-DP.";  
 RP Genes Dev. 13:666-674(1999).  
 RL [6]  
 RP FUNCTION: CAN STIMULATE E2F-DEPENDENT TRANSCRIPTION. BINDS DNA  
 RP COOPERATIVELY WITH E2F FAMILY MEMBERS THROUGH THE E2 RECOGNITION  
 RP SITE. TTTC/GCC, FOUND IN THE PROMOTER REGION OF A NUMBER OF  
 RP GENES WHOSE PRODUCTS ARE INVOLVED IN CELL CYCLE REGULATION OR IN  
 RP DNA REPLICATION. THE DP2/E2F COMPLEX FUNCTIONS IN THE CONTROL OF  
 RP CELL-CYCLE PROGRESSION FROM G1 TO S PHASE. THE E2F-1/DP COMPLEX  
 RP APPEARS TO REGULATE BOTH CELL PROLIFERATION AND APOPTOSIS.  
 RP -1- SUBUNIT: COMPONENT OF THE DP2/E2F TRANSCRIPTION FACTOR COMPLEX.  
 RP FORMS HETERODIMERS WITH E2F FAMILY MEMBERS. THE COMPLEX CAN  
 RP INTERACT WITH HYPOPHOSPHORYLATED RETINOBLASTOMA PROTEIN RB1 AND  
 RP RELATED PROTEINS (P107 AND P130) THAT INHIBIT THE E2F  
 RP TRANSACTIVATION DOMAIN. DURING THE CELL CYCLE, RB BECOMES  
 RP PHOSPHORYLATED IN MID-TO-LATE G1 PHASE, DETACHES FROM THE  
 RP DP2/E2F COMPLEX RENDERING E2F TRANSCRIPTIONALLY ACTIVE. VIRAL  
 RP ONCOPROTEINS, NOTABLY E1A, T-ANTIGEN AND SV40 E7, ARE CAPABLE OF

CC SEQUESTERING RB PROTEIN, THUS RELEASING THE ACTIVE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS: ALPHA/49 KDA (SHOWN  
 CC HERE), BETA/43 KDA, GAMMA AND DELTA/48 KDA. MAY BE PRODUCED BY  
 CC ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: HIGH LEVELS IN HEART AND SKELETAL MUSCLE. ALSO  
 CC FOUND IN PLACENTA, KIDNEY, BRAIN, LUNG AND LIVER. THE PRESENCE AS  
 CC WELL AS THE ABUNDANCE OF THE DIFFERENT TRANSCRIPTS APPEAR TO  
 CC VARY SIGNIFICANTLY IN DIFFERENT TISSUES AND CELL LINES.  
 CC -1- PTM: PHOSPHORYLATED.  
 CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use, by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.1sb-std.ch/announce/>  
 CC or send an email to [license@1sb-std.ch](mailto:license@1sb-std.ch)).  
 CC -----  
 CC EMBL: L40386; AAC63016.1;  
 CC EMBL: U18422; AAB60378.1;  
 CC EMBL: U75488; AAB37321.1;  
 CC EMBL: U75117; AAC50642.1;  
 CC PDB: 1CP7; 02-APR-99.  
 CC TRANSFAC: T03000;  
 CC MIM: 602160;  
 CC InterPro: IPR003316; E2F\_TDP.  
 CC Pfam: PF002319; E2F\_TDP.1.  
 CC Transcription regulation; DNA-binding; Activator; Nuclear protein;  
 CC Phosphorylation; Cell cycle; Alternative splicing; 3D-structure.  
 CC DNA\_BIND 129 210  
 CC FT 176 210 DEF BOX  
 CC FT 219 292  
 CC FT 229 261  
 CC FT 274 330  
 CC FT 432 446 ASP/GLU-RICH (ACIDIC) (NCB DOMAIN).  
 CC FT 432 446 PHOSPHORYLATION (BY CDK2) (POTENTIAL).  
 CC FT 42 42 PHOSPHORYLATION (BY CDK2) (POTENTIAL).  
 CC FT 1 61 MISSING (IN ISOFORM BETA, ISOFORM GAMMA  
 CC AND ISOFORM DELTA).  
 CC FT 103 118 MISSING (IN ISOFORM BETA AND ISOFORM  
 CC GAMMA).  
 CC FT 173 173 S -> SQ (IN ISOFORM GAMMA).  
 CC FT 64 64 I -> T.  
 CC FT VARIANT  
 CC SEQUENCE 446 AA; 49236 MW; 19A6C85BADE1DFE1 CRC64;  
 CC FT 50

Query Match 100.0%; Score 186; DB 1; Length 446;  
 Best Local Similarity 100.0%; Pred. No. 1e-18;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KNIRRYVDALVLMANNISKEKKEIKWGLPTNSA 37  
 |||||||||||||||||||||||||||||||||||  
 DB 178 KNIRRYVDALVLMANNISKEKKEIKWGLPTNSA 214

RESULT 4  
 TDP\_DROME STANDARD; PRT; 377 AA.  
 AC Q24318;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Transcription factor dp.  
 GN DP.  
 OS Drosophila melanogaster (fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Phyllophaga; Drosophilidae; Drosophillinae.  
 NC NB1\_Tasid-7217;  
 RX SOURCE FROM N.A. (DELTA ISOFORM), AND CHARACTERIZATION.  
 RX MEDLINE-95357935; PubMed-779537;  
 RX Hu C.L., Sukerberg L.R., Ngwu C., Harlow E., Lees J.A.;  
 RX "In vivo association of E2F and DP family proteins.";  
 RX Mol. Cell. Biol. 15:2536-2546(1995).  
 RL [2]  
 RP SEQUENCE FROM N.A. (DELTA ISOFORM).  
 RP TISSUE-Kidney; PubMed-7784053;  
 RP MEDLINE-9530370; PubMed-7784053;  
 RP Zhang Y., Chellappan S.P.;  
 RP Cloning and characterization of human DP2, a novel dimerization  
 RP partner of E2F.";  
 RP Oncogene 10:7085-7093(1995).  
 RL [3]  
 RP SEQUENCE OF 1-102 FROM N.A. (ALPHA ISOFORM).  
 RP Zhang Y., Chellappan S.P.;  
 RP "Transcriptional activation and expression of DP transcription  
 RP factor during cell cycle and TPA-induced U937 differentiation.";  
 RP Submitted (DEC-1996) to the EMBL/GenBank/DBS databases.  
 RL [4]  
 RP SEQUENCE OF 63-79 FROM N.A. (DELTA ISOFORM), AND ALTERNATIVE SPLICING.  
 RP TISSUE-NerveTissue;  
 RP MEDLINE-96353862; PubMed-8755520;  
 RP Rogers K.T., Higgins P.D.R., Milla M.M., Phillips R.S., Horowitz J.M.;  
 RP "DP-2, a heterodimeric partner of E2F: identification and  
 RP characterization of DP-2 proteins expressed in vivo.";  
 RP Proc. Natl. Acad. Sci. U.S.A. 93:7594-7599(1996).  
 RL [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 121-215.  
 RP MEDLINE-99190701; PubMed-10090723;  
 RP Zheng M., Freeman E., Pabo C.O., Pavletich N.P.;  
 RP "Structural basis of DNA recognition by the heterodimeric cell cycle  
 RP transcription factor E2F-DP.";  
 RP Genes Dev. 13:666-674(1999).  
 RL [6]  
 RP FUNCTION: CAN STIMULATE E2F-DEPENDENT TRANSCRIPTION. BINDS DNA  
 RP COOPERATIVELY WITH E2F FAMILY MEMBERS THROUGH THE E2 RECOGNITION  
 RP SITE. TTTC/GCC, FOUND IN THE PROMOTER REGION OF A NUMBER OF  
 RP GENES WHOSE PRODUCTS ARE INVOLVED IN CELL CYCLE REGULATION OR IN  
 RP DNA REPLICATION. THE DP2/E2F COMPLEX FUNCTIONS IN THE CONTROL OF  
 RP CELL-CYCLE PROGRESSION FROM G1 TO S PHASE. THE E2F-1/DP COMPLEX  
 RP APPEARS TO REGULATE BOTH CELL PROLIFERATION AND APOPTOSIS.  
 RP -1- SUBUNIT: COMPONENT OF THE DP2/E2F TRANSCRIPTION FACTOR COMPLEX.  
 RP FORMS HETERODIMERS WITH E2F FAMILY MEMBERS. THE COMPLEX CAN  
 RP INTERACT WITH HYPOPHOSPHORYLATED RETINOBLASTOMA PROTEIN RB1 AND  
 RP RELATED PROTEINS (P107 AND P130) THAT INHIBIT THE E2F  
 RP TRANSACTIVATION DOMAIN. DURING THE CELL CYCLE, RB BECOMES  
 RP PHOSPHORYLATED IN MID-TO-LATE G1 PHASE, DETACHES FROM THE  
 RP DP2/E2F COMPLEX RENDERING E2F TRANSCRIPTIONALLY ACTIVE. VIRAL  
 RP ONCOPROTEINS, NOTABLY E1A, T-ANTIGEN AND SV40 E7, ARE CAPABLE OF

RP SEQUENCE FROM N.A.  
 RC TISSUE-Dye Imaginal disk;  
 RA MEDLINE-94294381; PubMed-8022767;  
 RA Dylachet B., Brook A., Dembakl M., Tenush L., Dyson N.;  
 RA "DNA-binding and trans-activation properties of Drosophila E2F and DP  
 RT proteins";  
 RT Proc. Natl. Acad. Sci. U.S.A. 91:6359-6363(1994).  
 CC -1- FUNCTION: CAN STIMULATE E2F-DEPENDENT TRANSCRIPTION.  
 CC -1- SUBUNIT: HETEROOLIGOMER OF E2F AND DP COOPERATE TO GIVE SEQUENCE-  
 CC SPECIFIC DNA BINDING AND OPTIMAL TRANS-ACTIVATION.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potentially).  
 CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.  
 CC .....  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: X79708; CA556147.2; ALT\_INT.  
 DR HSP: Q1188; ICF7.  
 DR Flybase: FBgn0011763; DP.  
 DR Interpro: IPR003116; E2F\_TDP.  
 DR Pfam: PF02319; E2F\_TDP; 1.  
 DR Transcription regulation; DNA-binding; Nuclear protein.  
 FT DOMAIN 81 88 POLY-SER.  
 FT SEQUENCE 377 AA; 42683 MW; 868F180243C3DC CRC64.

Query Match 89.23; Score 166; DB 1; Length 377;  
 Best Local Similarity 86.11; Pred. No. 5.6e-16;  
 Matches 31; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 KNIRRVYDALVLAANNISKEKIKNGLP7NS 36  
 DB 145 KNIRRVYDALVLAANNISKEKIKNGLP7NS 180

RESULT 5  
 TOPH\_CAREL STANDARD; PRT; 598 AA.  
 AC 023703;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Transcription factor DP homolog.  
 GN T2367.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Plectiderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Baiton K.;  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 RM (2)  
 RM REVISIONS.  
 RC STRAIN-BRISTOL N2;  
 RA Dublin R.;  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potentially).  
 CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.  
 CC .....  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: I68319; CA932699.1;  
 DR HSP: Q14188; ICF7.  
 DR Wormpep: T2367.1; CE03702.  
 DR Interpro: IPR003116; E2F\_TDP.  
 DR Pfam: PF02319; E2F\_TDP; 1.  
 DR Hypothetical protein; Transcription regulation; DNA-binding;  
 RW Nuclear protein.  
 SO SEQUENCE 598 AA; 67930 MW; A79692087F3D3DAE CRC64;

Query Match 82.38; Score 153; DB 1; Length 598;  
 Best Local Similarity 75.75; Pred. No. 6.3e-14;  
 Matches 28; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

OY 1 KNIRRVYDALVLAANNISKEKIKNGLP7NS 37  
 DB 122 KNIRRVYDALVLAANNISKEKIKNGLP7NS 158

RESULT 6  
 TOP2\_MOUSE STANDARD; PRT; 446 AA.  
 AC 06163;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Transcription factor DP-2 (E2F dimerization partner 2) (DP-3).  
 GN TPDP2 OR DP2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.  
 RC TISSUE-brain, and kidney;  
 RA MEDLINE-96038871; PubMed-7478568;  
 RA O'Rourke D.E., De La Lanza S., La Thangue N.B.;  
 RA "A new member of the DP family, DP-3, with distinct protein products  
 RT suggests a regulatory role for alternative splicing in the cell cycle  
 RT transcription factor DP2/2P2";  
 RT Oncogene 11:1437-1446(1995).  
 RL -1- FUNCTION: CAN STIMULATE E2F-DEPENDENT TRANSCRIPTION. BINDS DNA  
 CC COOPERATIVELY WITH E2F FAMILY MEMBERS THROUGH THE E2 RECOGNITION  
 CC SITE. TTCC/GCCG, FOUND IN THE PROMOTER REGION OF A NUMBER OF  
 CC GENES WHOSE PRODUCTS ARE INVOLVED IN CELL CYCLE REGULATION OR IN  
 CC DNA REPLICATION. THE DP2/E2F COMPLEX FUNCTIONS IN THE CONTROL OF  
 CC CELL-CYCLE PROGRESSION FROM G1 TO S PHASE. THE E2F-1/DP COMPLEX  
 CC APPEARS TO MEDiate BOTH CELL PROLIFERATION AND APOPTOSIS.  
 CC -1- SUBUNIT: COMPONENT OF THE DP2/E2F TRANSCRIPTION FACTOR COMPLEX.  
 CC FORMS HETEROOLIGOMERS WITH E2F FAMILY MEMBERS. THE COMPLEX CAN  
 CC INTERACT WITH HYPOPHOSPHORYLATED RETINOBLASTOMA PROTEIN RB1 AND  
 CC RELATED PROTEINS (P107 AND P130) THAT INHIBIT THE E2F  
 CC TRANSCRIPTIONAL DOMAIN. DURING THE CELL CYCLE, RB BECOMES  
 CC PHOSPHORYLATED IN MANY TO-LATE G1 PHASE, DETACHES FROM THE  
 CC DP2/E2F COMPLEX, RENDERING E2F TRANSCRIPTIONALLY ACTIVE. VITAL  
 CC ONCOPROTEINS, NOTABLY E1A, T-ANTIGEN AND HPV E7, ARE CAPABLE OF  
 CC SUBVERTING RB PROTEIN, THUS RELASING THE ACTIVE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOPROTEINS, ALPHA (SHOWN HERE),  
 CC BETA, GAMMA AND DELTA, ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED. HIGHEST  
 CC LEVELS IN SPLEEN AND HEART.  
 CC -1- PTM: PHOSPHORYLATION BY E2F-1-BOUND CYCLIN A-COR2, IN THE S PHASE,  
 CC INHIBITS E2F-MEDIATED DNA BINDING AND TRANSCRIPTION.  
 CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.  
 CC .....  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>

```

Query Match      Score 98.5: DB 1: Length 446;
Best Local Similarity 67.66: Pred. No. 2,3e-05;
Matches 23: Conservative 1: Mismatches 1: Indels 9: Caps 1:

Oy 1 KNIRRYVALNLYAAMNITSKKKIKIGIDPT 34
      :|||||:|||||:|||||:|||||:
Db 178 ENIRRYVALNLYAAMNITS-----SLPT 202

```

CC	entire requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@lsdb.sib.ch">license@lsdb.sib.ch</a> ).	
CC		
DR	ENBL: NP032131; M053521.1; -	
DR	HSSP: O16254; 1CF7.	
DR	TRANSPAC: T03250; -	
DR	MCD: MGI:1354159; x3f6.	
DR	InterPro: IPR003316; x3f.TDP.	
DR	Pfam: PF02319; x3f.TDP.1	
K4	Transcription regulation; DNA-binding; Repressor; Nuclear protein; Cell cycle.	
K4		
K4		
F3	DOMAIN	1 62
F3	DNA_BIND	50 126
F3		129
F7	DOMAIN	95 122
F7		129
F7	DOMAIN	130 222
F7		164
F7	DOMAIN	143 272
F7		
FE	SEQUENCE	272 AA; 30882 MW; CF425510EADBBDB8 CRTC4;

```

RESULT 8
E2F6_HUMAN STANDARD; PRT; 281 AA.
ID E2F6_HUMAN
IC 075461: 060544;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor E2F6 (E2F-6).
GN E2F6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OC NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98356133; PubMed-9689056;
RX Gaudat S., Wood J.O., Livingston D.M.;
RT 'Unusual proliferation arrest and transcriptional control properties
RT of a newly discovered E2F family member, E2F-6.';
RL Proc. Natl. Acad. Sci. U.S.A. 95:9190-9195(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-98356235; PubMed-9704927;
RX Carveright P., Muller W., Wagner C., Holm K., Mellin K.;
RT 'E2F-6: a novel member of the E2F family is an inhibitor of E2F-
RT dependent transcription.';
RL Oncogene 17:611-623(1998).
RN [3]
RP SEQUENCE OF 7-281 FROM N.A.
RX TISSUE-Petal brain;
RX MEDLINE-98159460; PubMed-9501179;
RX Trismarchi J.M., Patchild B., Verone R., Moberg K., Andon N.;
RT 'Leea J.A., member of the E2F family that can behave as a
RT transcriptional repressor.';
RL Proc. Natl. Acad. Sci. U.S.A. 95:2850-2855(1998).
CC -1. FUNCTION: INHIBITOR OF E2F-DEPENDENT TRANSCRIPTION. BINDS DNA
CC SPECIFICALLY WITH DP PROTEINS THROUGH THE 53 RECOGNITION SITE,
CC TTTCGGCC. HAS A PRESENCE FOR THE TTTCGGC E2F RECOGNITION
CC SITE. E2F-6 LACKS THE TRANSCRIPTIONAL ACTIVATE AND POCKET
CC PROTEIN BINDING DOMAINS. APPEARS TO REGULATE A SUBSET OF E2F-
CC DEPENDENT GENES WHOSE PRODUCTS ARE REQUIRED FOR ENTRY INTO THE
CC CELL CYCLE BUT NOT FOR NORMAL CELL CYCLE PROGRESSION.

```

```
CC OVEREXPRESSION DELAYS THE EXIT OF CELLS FROM THE S-PHASE.
CC -1- SUBUNIT: COMPONENT OF THE DRP1/E2F TRANSCRIPTION FACTOR COMPLEX.
CC FORMS HETEROOLIGOMERS WITH DP FAMILY MEMBERS.
CC CC
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC CC
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED. HIGHEST
CC LEVELS IN PLACENTA, SKELETAL MUSCLE, HEART, OVARY, KIDNEY, SMALL
CC INTESTINE AND SPLEEN.
CC CC
CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright . It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC CC
CC EMBL: AF059292; AAC31426.1; -.
CC DR EMBL: AF041381; AAC14694.1; -.
CC DR HSSP: Q16254; ICP7.
CC DR MIM: 602944; -.
CC DR INTERPRO: IPR003316; E2F-TDP.
CC DR Pfam: PF00219; E2F_TDP_1.
CC KW Transcription regulation; DNA-binding; Repressor; Nuclear protein; Cell cycle.
CC MW 50
CC FT DNA_BIND 50 129 POTENTIAL.
CC FT DOMAIN 95 129 DEF BOX.
CC FT DOMAIN 130 222 DIMERIZATION (POTENTIAL).
CC FT DOMAIN 143 164 LENTICINE-RIPPER.
CC FT DOMAIN 173 281 TRANSCRIPTION REPRESSION DOMAIN.
CC FT DOMAIN 265 268 POLY-GLU.
CC FT MUTAGEN 68 L-27E: REDUCTION IN REPRESSOR ACTIVITY, LITTLE EFFECT ON S-PHASE ENTRY.
CC FT CONFLICT 7 8 AR -> HE (IN REF. 3).
CC FT CONFLICT 220 220 I -> V (IN REF. 3).
CC FT CONFLICT 229 229 G -> E (IN REF. 3).
CC SQ SEQUENCE 281 AA; 31844 MW; 539E049C13AD3508 CRC64;
CC
QY Query Match 39.0% Score 72.5; DB 1; Length 281;
DB Best Local Similarity 48.3%; Pred.No. 0.0066;
Matches 14; Conservative 7; Mismatches 7; Indels 1; Gaps 1;
Db 100 KRATVDTNVNLCIDIVLKSKMHRWVG 126
    :||||| ||| :::: | : | ||||
    4 RRRVYDALNLVANNIISRE-KREIKMIC 31
    100 KRATVDTNVNLCIDIVLKSKMHRWVG 126

```

RESULT: 9

ID	E2F1_MOUSE	STANDARD:	PRT:	430 AA.
NC	061501;			
D7	01-MOV-1997 (Rel. 35, Created)			
D7	01-MOV-1997 (Rel. 35, Last sequence update)			
D7	30-MAY-2000 (Rel. 39, Last annotation update)			
D7	Transcription factor E2f1 (E2F-1).			
GN	E2f1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCB1_TaxId=10090;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SWISS ALBINO;			
RA	MEDLINE=44188985; PubMed=8114719;			
RA	L.I.V., Slansky J.E., Myers D.J., Drinkwater M.R., Kaelin W.G.,			
PA	Farnham P.J.;			
RF	"Cloning, chromosomal location, and characterization of mouse E2f1."			
RL	Mol. Cell. Biol. 14:1861-1869(1994).			
RN	(2)			
RD	DEVELOPMENTAL EXPRESSION.			
RX	MEDLINE=98075478; PubMed=9376316;			
RX	Daganzo L., Fry C.J., Bartley S.M., Farnham P., Gollic B.L.,			

RA PHILIP R. N.; "Expression patterns of the E2F family of transcription factors during  
R2 mouse nervous system development."; *Neurosci. Dev.* 66:13-25(1997).  
R3 Mech. Dev. 66:13-25(1997).  
R4 [1].  
R5 FUNCTION IN APOPTOSIS.  
R6 MEDLINE-9837455; PubMed-9674698;  
R7 Holmberg K., Hellen K., Scheked M., Kristensen O.;  
R8 "E2F-1-induced p53-independent apoptosis in transgenic mice.";  
R9 Oncogene 17:143-155(1998).  
R10  
R11 -1- FUNCTION: TRANSCRIPTION ACTIVATOR THAT BINDS DNA COOPERATIVELY  
R12 WITH DP PROTEINS THROUGH THE E2 RECOGNITION SITE, TTTC/GCGC.  
R13 FOUND IN THE PROMOTER REGION OF A NUMBER OF GENES WHOSE PRODUCTS  
R14 ARE INVOLVED IN CELL CYCLE REGULATION OR IN DNA REPLICATION. THE  
R15 ORF1/E2F COMPLEX FUNCTIONS IN THE CONTROL OF CELL-CYCLE  
R16 PROGRESSION FROM G1 TO S PHASE. E2F-1 BINDS PREFERENTIALLY RAI  
R17 PROTEIN. IN A CELL-CYCLE DEPENDENT MANNER, IT CAN MEDIANE BOTH  
R18 CELL PROLIFERATION AND p53-DEPENDENT APOPTOSIS.  
R19  
R20 -1- SUBUNIT COMPONENT OF THE DRP1/E2F TRANSCRIPTION FACTOR COMPLEX.  
R21 FORMS HETERODIMERS WITH DP FAMILY MEMBERS. THE E2F-1 COMPLEX BINDS  
R22 SPECIFICALLY HYPOPHOSPHORYLATED RETINOBLASTOMA PROTEIN RB1. DURING  
R23 THE CELL CYCLE, RB1 BECOMES PHOSPHORYLATED IN MID-TO-LATE G1  
R24 PHASE, DETACHES FROM THE DRP1/E2F COMPLEX, RENAMING E2F  
R25 TRANSCRIPTIONALLY ACTIVE. VIRAL ONCOPROTEINS, NOTABLY E1A, T-  
R26 ANTIGEN AND HPV E7, ARE CAPABLE OF SEQUESTERING RB PROTEIN, THUS  
R27 RELEASING THE ACTIVE COMPLEX.  
R28  
R29 -1- SUBCELLULAR LOCATION: Nucleus.  
R30  
R31 -1- DEVELOPMENTAL STAGE: IN THE DEVELOPING NERVOUS SYSTEM, FIRST  
R32 DETECTED IN THE NEURAL TUBE AT DAY 9.5 DPC. BY DAY 10.5, LEVELS  
R33 INCREASE THROUGHOUT THE BRAIN, WITH HIGHEST LEVELS IN THE  
R34 HINDRAIN AND IN THE SPINAL CORD, EXPRESSED ONLY IN THE ROSTRAL  
R35 HALF. BY DAY 11.5, EXPRESSION POUND THROUGHOUT THE BRAIN AND  
R36 SPINAL CORD. FROM DAY 12.5, EXPRESSION RESTRICTED TO THE  
R37 VENTRICULAR REGIONS OF THE BRAIN, PEAKS AT DAY 13.5 AND DECLINES  
R38 THEREAFTER. ONLY WEAK EXPRESSION IN THE DEVELOPING SPINAL CORD  
R39 FROM DAY 11.5-16.5. IN THE DEVELOPING RETINA, EXPRESSION IS  
R40 CONFINED TO THE UNDIFFERENTIATED RETINOBLASTIC CELL LAYER IN  
R41 OTHER DEVELOPING TISSUES, E2F-1 IS EXPRESSED IN KIDNEY, LUNG,  
R42 LIVER HEPATOCYTES, HEART AND THYMUS. HIGHEST LEVELS IN LIVER.  
R43  
R44 -1- PTV: PHOSPHORYLATED BY CDK2 AND CYCLIN A-CDK2 IN THE S-PHASE (BY  
R45 SIMILARITY).  
R46  
R47 -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.  
R48  
R49 THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION  
R50 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -  
R51 CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON LEA-  
R52 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY  
R53 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL  
R54 CC ENTITIES REQUIRES A LICENSE AGREEMENT (See <http://www.lsb.slb.ch/announce/>  
R55 CC or send an email to [license@lsb.slb.ch](mailto:license@lsb.slb.ch)).  
R56  
R57 EMBL: L21973; AAC83217.1;  
R58 DR HSP: Q16754; ICF;  
R59 DR TRANSFAC: T00219;  
R60 DR TRANSFAC: T01563;  
R61 DR MCD: M01101941; E2F1.  
R62 DR INTERPRO: IPR003316; E2F\_RDP.  
R63 DR Pfam: PF03319; E2F\_RDP; 1.  
R64 DR Transcription regulation; Activator; DNA-binding; Nuclear protein;  
R65 DR Phosphorylation; Cell cycle; Apoptosis.  
R66 CYCLIN A/CDK2 BINDING (POTENTIAL).  
R67 DNA BIND 62 103  
R68 DR DNA BIND 105 189  
R69 DR DOMAIN 148 169  
R70 DR DOMAIN 153 189  
R71 DR DOMAIN 190 279  
R72 DR DOMAIN 361 430  
R73 DR DOMAIN 402 419  
R74  
R75 SEQUENCE 430 AA; 46323 MW; CSDFI8ND3BADPFA CRC64;  
R76  
R77  
R78  
R79  
R80

Best local similarity: 43.8%; Pred. No.: 0.014; Matches: 11; Conservative: 8; Mismatches: 9; Indels: 1; Gaps: 1

07 4 RAYTOLANTLANNIISKE-KKICIKGLPT 34  
          |||         |||         |||  
Db 159 KRRIDITNVLGGIOILKKSKKHNIOWLSHF 190

RESULT 10  
ID E2F1\_HUMAN STANDARD: PRT, 437 AA.  
AC 001094: Q92768: Q13143:  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DI 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Transcription factor E2F1 (E2F-1) (retinoblastoma binding protein 3)  
DE (RBP-3) (PRB-binding protein E2F-1) (PRB3) (Retinoblastoma-associated  
DB protein 1) (RBP-1).  
GN E2F1 OR RBP3  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_Taxid=9606;  
CN 1

RX MEDLINE FROM N.A. PubMed=1638634;  
RX MEDLINE-92146721; PubMed=1638634;  
RX Kneilo W.G., Jr, Krek W., Sellers W.R., Decaprio J.A., Ajchenbaum P.,  
RX Puche C.S., Chittenden T., Li Y., Farnham P.J., Blaser M.A.,  
RA Livingston D.M., Plomington E.K., Farnham P.J., Blaser M.A.,  
RW "Expression cloning of a cDNA encoding a retinoblastoma-binding  
RW protein with E2F-like properties.";  
RW Cell 70:337-350(1992).  
RN 12  
RP SEQUENCE FROM N.A.  
RL MEDLINE-92146721; PubMed=1638635;  
RL Kneilo W.G., Jr, Krek W., Sellers W.R., Decaprio J.A., Ajchenbaum P.,  
RN Puche C.S., Chittenden T., Li Y., Farnham P.J., Blaser M.A.,  
RL Livingston D.M., Plomington E.K., Farnham P.J., Blaser M.A.,  
RL "Expression cloning of a cDNA encoding a retinoblastoma-binding  
RL protein with E2F-like properties.";  
RL Cell 70:337-354(1992).

RX MEDLINE FROM N.A. PubMed=1448092;  
RX MEDLINE-93078763; PubMed=1448092;  
RX Shan B., Zhu X., Chen P.L., Dutfee T., Yang X., Sharp D., Lee W.H.;  
RW Molecular cloning of cellular genes encoding  
RW retinoblastoma-associated proteins: identification of a gene with  
RW properties of the transcription factor E2F.";  
RW Mol. Cell. Biol. 12:15620-15631(1992).  
RN 14  
RP SEQUENCE FROM N.A.  
RL MEDLINE-97082961; PubMed=8964493;  
RL Newman E., Sellers W.R.S., McNeil J.A., Lawrence J.B.,  
RA Kneilo W.G., Jr,  
RW "Structure and partial genomic sequence of the human E2F1 gene.";  
RW Gene 171:163-169(1996).

RX MEDLINE FROM N.A.  
RX DeDuke P., Matthews L.R., Ashurst J., Burton J., Gilbert J.G.R.,  
RX Jones M., Steavides G., Almeida J.P., Babbage A.K., Baggaley C.L.,  
RX Bailey J., Barlow K.F., Bates K.N., Beard L.N., Beare D.M.,  
RX Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RX Buck D., Burrill W., Butler A.P., Carder C., Carter N.P.,  
RX Chapman J.C., Champ M., Clark G., Clark L.N., Clark S.V., Clee C.M.,  
RX Clegg S., Cobley V.E., Collier R.B., Connor F., Corby M.R.,  
RX Coulson A., Coville G.J., Dedman R., Daniel F., Dunn M.,  
RX Ellinghaus A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RX Graham D.V., Griffiths C., Griffiths M.M.D., Gilliland R., Hall R.B.,  
RX Hammond E., Hiley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RX Huckle E., Hunt A.R., Hunt S.E., Jechsch K., Johnson C.W., Johnson D.,  
RX Kay M.P., Kimbroley A.M., King T., Knights A., Laird G.K., Taylor S.,  
RX Leharzello M.B., Leverish N., Lloyd C., Lloyd D.M., Lovell J.D.,  
RX March V.L., Martin S.L., McCormacke L.J., Melay K., McNulty A.A.,  
RX Milner K., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RX Oliver A., Parler A., Patel K., Pearce T.A.V., Peck A.I.,

RA PHILLIPS B.J.C.T., Preheltingem S.R., Plumb R.W., Ramsay N.  
RA Rice C.M., Ross M.T., Scott C.E., Sebra H.K., Shonkeen R., Sims S.  
RA Shue C.D., Smith M.L., Soderland C., Stevard C.A., Sulston J.B.,  
RA Stuenkel M., Symons R.M., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Treman A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead E.L., Whitaker P., Willey D.L., Williams L., Williams S.,  
RA Witting L., Wray F.W., Hubbard T., Barlow R.N., Bentley D.R., Beck  
RA Rogers J.; The DNA sequence and comparative analysis of human chromosome 20.  
RL Nature 418:665-671(2001).  
RU  
RV  
RW  
RX SEQUENCE OF 1-111 FROM M.A.  
RY MEDLINE-95047311; PubMed-7958936;  
RZ Johnson D.G., Ohtani K., Nevins J.R.;  
SA "Autoregulation control of E2F1 expression in response to positive  
SB and negative regulators of cell cycle progression";  
SC Gene Dev. 8:1514-1525(1994).  
SD  
SE  
SF PHOSPHORYLATION.  
SG MEDLINE-9510413; PubMed-7938933;  
SH Kitagawa M., Higashi H., Suzuki-Takahashi I., Sogawa K., Hanke S.K.  
SI Klagsberg M., Hignash H., Shiratori S., Arany Z., Keelin M.G. Jr.,  
SJ Reek M., Even M.E., Shirokhar S., Arany Z., Keelin M.G. Jr.,  
SK Phosphorylation of E2F-1 by cyclin A-cdk2.";  
SL Oncogene 10:229-236(1995).  
SM  
SN  
SO CYCLIN A/CDK2 BINDING DOMAIN.  
SP MEDLINE-94306513; PubMed-8033208;  
SQ Reek M., Even M.E., Shirokhar S., Arany Z., Keelin M.G. Jr.,  
SR Negative regulation of the growth-promoting transcription factor  
SS E2F-1 by a stably bound cyclin A-dependent protein kinase.";  
ST Cell 78:165-172(1994).  
SU  
SV  
SW  
SX [9]  
SY DIFFERENTIAL REGULATION BY CYCLIN/CDK2 KINASES.  
SZ MEDLINE-95047313; PubMed-7958936;  
TA Dynlacht B.D., Flores O., Lees J.A., Harlow E.;  
TB Differential regulation of E2F transactivation by cyclin/cdk2  
TC complexes.";  
TD Games Dev. 8:1772-1786(1994).  
TE  
TF  
TG  
TH  
TI REGULATION BY CYCLIN-DEPENDENT KINASES.  
TJ MEDLINE-97342624; PubMed-1939321;  
TK Dynlacht B.D., Nobury K., Lees J.A., Harlow E., Zhu L.;  
TL "Specific regulation of E2F family members by cyclin-dependent  
TM kinases.";  
TN Mol. Cell. Biol. 17:3867-3875(1997).  
TO  
TP  
TQ INHIBITION OF DNA-BINDING.  
TR MEDLINE-95050971; PubMed-7968976;  
TS Xu M., Sheppard K.-Y., Peng C.-Y., Yee A.S., Plyusce-Morne M.;  
TT "Cyclin A-cdk2 binds directly to E2F-1 and inhibits the DNA-binding  
TU activity of E2F-1/DP-1 by phosphorylation.";  
TV Mol. Cell. Biol. 14:8420-8431(1994).  
TW  
TX  
TY  
TZ  
UA  
UB  
UC  
UD  
UE  
UF  
UG  
UH  
UI  
UJ  
UK  
UL  
UM  
UN  
UU  
UV  
VW  
VX  
VY  
VZ  
WA  
WB  
WC  
WD  
WE  
WF  
WG  
WH  
WI  
WJ  
WK  
WL  
WM  
WN  
WO  
WP  
WQ  
WR  
WS  
WT  
WX  
WY  
WZ  
XA  
XB  
XC  
XD  
XE  
XF  
XG  
XH  
XI  
XJ  
XK  
XL  
XM  
XN  
XO  
XP  
XQ  
XR  
XS  
XT  
XU  
XV  
XW  
XX  
XY  
XZ  
YA  
YB  
YC  
YD  
YE  
YF  
YG  
YH  
YI  
YJ  
YK  
YL  
YM  
YN  
YO  
YP  
YQ  
YR  
YS  
YT  
YU  
YV  
YW  
YX  
YY  
YZ  
ZA  
ZB  
ZC  
ZD  
ZE  
ZF  
ZG  
ZH  
ZI  
ZJ  
ZK  
ZL  
ZM  
ZN  
ZO  
ZP  
ZQ  
ZR  
ZS  
ZT  
ZU  
ZV  
ZW  
ZX  
ZY  
ZZ

[illegible]

```

01 4 RRVYDQALNYLMMNNISRE-KKEIKNIG 31
02 11:11 11: :1:1:1:1:1:1
Db 166 KRIYDTINVLGGIOLIRKANNHNTOWNG 194

RESULT 13
E2F3_MOUSE STANDARD: PRT: 356 AA.
ID E2F3_MOUSE 035161;
DC 15-JUL-1998 (rel. 36, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DV 30-MAY-2000 (rel. 39, Last annotation update)
DE Transcription factor E2F3 (E2F-3) [Fragment].
GN E2F3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrates; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
PI 1;
RF SEQUENCE FROM N.A. STRAIN-SWISS ALBINO; TISSUE=fibroblast;
RX MEDLINE=98025478; PubMed=9376316;
RA Daigino L., Fry C.J., Bartley S.M., Farnham P., Galle B.L.,
RA Phillips R.A.;
RA Phillips R.A.:
RA Expression patterns of the E2F family of transcription factors
RT during mouse nervous system development.?
RL Mech. Dev. 66:13-25(1997).
RN (2)
RP DEVELOPMENTAL EXPRESSION:
RX MEDLINE=97294022; PubMed=9193906;
RA Daigino L., Fry C.J., Bartley S.M., Farnham P., Galle B.L.,
RA Phillips R.A.;
RA Phillips R.A.:
RA Expression patterns of the E2F family of transcription factors
RT during murine epithelial development.?
RL Cell Growth Differ. 8:553-563(1997).
RN -1
RF FUNCTION: TRANSCRIPTION ACTIVATOR THAT BINDS DNA COOPERATIVELY
WITH DP PROTEINS THROUGH THE E2 RECOGNITION SITE, TTTC/GGCC,
CC FOUND IN THE PROMOTER REGION OF A NUMBER OF GENES WHOSE PRODUCTS
CC ARE INVOLVED IN CELL-CYCLE REGULATION OR IN DNA REPLICATION. THE
CC DREF1/E2F COMPLEX FUNCTIONS IN THE CONTROL OF CELL-CYCLE
CC PROGRESSION FROM G1 TO S PHASE. E2F-3 BINDS SPECIFICALLY TO RB1
CC PROTEIN. IN A CELL-CYCLE DEPENDENT MANNER.
CC -1 SUBUNIT COMPONENT OF THE DREF1/E2F TRANSCRIPTION FACTOR COMPLEX.
CC BINDS COOPERATIVELY WITH CP-1 TO E2F SITES. INTERACTS WITH
CC RETINOBLASTOMA PROTEIN RB1 AND RELATED PROTEINS (SUCH AS p107)
CC THAT INHIBIT THE E2F TRANSACTIVATION DOMAIN.
CC -1 SUBCELLULAR LOCATION: Nuclear.
CC -1 DEVELOPMENTAL STAGE: IN THE DEVELOPING NERVOUS SYSTEM, HIGH LEVELS
CC EXPRESSED IN BOTH VENTRAL AND DORSAL REGIONS OF THE SPINAL CORD
CC FROM 13.5 DPC. ALSO EXPRESSED IN DORSAL ROOT AND CRANIAL GANGLIA
CC IN 11.5-18.5 DPC EMBRYOS. ONLY LOW LEVELS OF EXPRESSION IN
CC DEVELOPING BRAIN. IN THE DEVELOPING RETINA (15.5 DPC), EXPRESSION
CC OF E2F-3 IS LOCALIZED TO THE GANGLION CELL LAYER. IN OTHER
CC DEVELOPING TISSUES,EXPRESSED IN LIVER, LUNG AND HEART. MEAN
CC EXPRESSION IN DEVELOPING KIDNEY AND SKELETAL MUSCLE. ABSENT FROM
CC THE DEVELOPING CHOROID PLEXUS, THYMUS AND DEVELOPING SKIN. LOW
CC LEVELS OF EXPRESSION IN THE DEVELOPING INTESTINAL EPITHELIUM AND
CC MESENCHYME IN 13.5-18.5 DPC EMBRYOS.
CC -1 SIMILARITY: BELONGS TO THE E2F/DP FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC
CC EMBL: AF015948; BAB1671.1;
CC HSPF_016354; ICF7.
CC TRANSFAC: T02947;
CC MOP: MOF1:1096340; EAT3.

```



DR InterPro: IPR003316; E2F\_TDP.  
 DR Pfam: PF02319; E2F\_TDP; 1.  
 KW Transcription regulation; Activator; DNA-binding; Nuclear protein;  
 KW Cell cycle.  
 FT NOVELTER 1 1  
 FT DOMAIN 1 44 CYCLIN A/CDK2 BINDING (POTENTIAL).  
 FT DNA\_BIND 46 136 POTENTIAL.  
 FT DOMAIN 95 116 LEUCINE-ZIPPER.  
 FT DOMAIN 100 136 DBF BOX.  
 FT DOMAIN 137 228 DIMERIZATION (POTENTIAL).  
 FT DOMAIN 282 356 TRANSACTIVATION (POTENTIAL).  
 FT DOMAIN 323 340 P130 PROTEIN ASSOCIATION (POTENTIAL).  
 FT SEQUENCE 356 AA; 35007 MW; 07BFC4E0A0C8E77 CRC64.

Query Match 36.8%; Score 68.5; DB 1; Length 356;  
 Best Local Similarity 41.4%; Pred. No. 0.031;  
 Matches 12; Conservative 9; Mismatches 7; Indels 1; Gaps 1;

OY 4 RRAYVDALVLMAMWISKE-KKEIKWIG 31  
 DB 106 KRRIIDITWVLCGICIKKSKRWVWNG 134

## RESULT 14

E2F3\_HUMAN STANDARD; PRT; 465 AA.

ID E2F3\_HUMAN 020716; Q15000;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 35, Last sequence update)  
 DE Transcription factor E2F3 (E2F-3).  
 GN E2F3 OR KIA00075.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94067143; PubMed-8246996;  
 RA Lees J.A., Saito M., Vidal M., Valantine M., Look T., Harlow E.,  
 DYON M., Hella K.;  
 RT "The retinoblastoma protein binds to a family of E2F transcription  
 RT factors.";  
 RT Mol. Cell. Biol. 13:7813-7825(1993).

RL [2]  
 RP SEQUENCE OF 292-465 FROM N.A.  
 RX MEDLINE-96051198; PubMed-7584044;  
 RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,  
 RA Bell N., Kawarabayashi Y., Ishikawa K.-I., Taddei S.;  
 RT "Prediction of the coding sequences of unidentified human genes. II.  
 RT Analysis of cDNA clones from human cell line KG-1.";  
 RT DNA Res. 1:223-229(1994).

CC -1- FUNCTION: TRANSCRIPTION ACTIVATOR THAT BINDS DNA COOPERATIVELY  
 WITH DP PROTEIN THROUGH THE E2 RECOGNITION SITE, TTTC/GCGC,  
 FOUND IN THE PROMOTER REGION OF A NUMBER OF GENES WHOSE PRODUCTS  
 ARE INVOLVED IN CELL CYCLE REGULATION OR IN DNA REPLICATION. THE  
 DBP/E2F COMPLEX FUNCTIONS IN THE CONTROL OF CELL-CYCLE  
 PROGRESSION FROM G1 TO S PHASE. E2F-3 BINDS SPECIFICALLY TO RB1  
 PROTEIN, IN A CELL-CYCLE DEPENDENT MANNER.  
 CC -1- SUBUNIT: COMPONENT OF THE DBP1/E2F TRANSCRIPTION FACTOR COMPLEX.  
 CC BINDS COOPERATIVELY WITH DB-1 TO E2F SITES. INTERACTS WITH  
 CC RETINOBLASTOMA PROTEIN RB1 AND RELATED PROTEINS (SUCH AS P107)  
 CC THAT INHIBIT THE E2F TRANSACTIVATION DOMAIN.  
 CC -1- SUBCELLULAR LOCATION: Nucleus.  
 CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.sdb.sdb.ch/announce/>  
 CC or send an email to [license@sdb.sdb.ch](mailto:license@sdb.sdb.ch)).

DR EMBL: Y10479; CAAT71504.1;  
 DR EMBL: D38550; BAA07553.1;  
 DR HSP: 016284; ICF7.  
 DR MIM: 600427.

DR InterPro: IPR003316; E2F\_TDP.  
 DR Pfam: PF02319; E2F\_TDP; 1.  
 KW Transcription regulation; Activator; DNA-binding; Nuclear protein;  
 KW Cell cycle.  
 FT NOVELTER 26 31  
 FT DOMAIN 45 53 POLY-ALA.  
 FT DOMAIN 101 133 CYCLIN A/CDK2 BINDING (POTENTIAL).  
 FT DOMAIN 120 129 POLY-GLY.  
 FT DNA\_BIND 355 245 POTENTIAL.  
 FT DOMAIN 204 225 LEUCINE-ZIPPER.  
 FT DOMAIN 209 245 DBF BOX.  
 FT DOMAIN 246 337 DIMERIZATION (POTENTIAL).  
 FT DOMAIN 391 465 TRANSACTIVATION (POTENTIAL).  
 FT DOMAIN 432 449 RETINOBLASTOMA PROTEIN ASSOCIATION  
 (POTENTIAL).  
 FT SEQUENCE 465 AA; 49161 MW; 461365842CA99EC CRC64;

Query Match 36.8%; Score 68.5; DB 1; Length 465;  
 Best Local Similarity 41.4%; Pred. No. 0.041;  
 Matches 12; Conservative 9; Mismatches 7; Indels 1; Gaps 1;

OY 4 RRAYVDALVLMAMWISKE-KKEIKWIG 31  
 DB 215 KRRIIDITWVLCGICIKKSKRWVWNG 243

## RESULT 15

E2F\_DROME STANDARD; PRT; 805 AA.

ID E2F\_DROME 027368; 077035;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Transcription factor E2F (DEZF).  
 GN E2F OR E2F1 OR CG6376.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;  
 OC Euryptera; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;

RP SEQUENCE FROM N.A.  
 RX MEDLINE-94294381; PubMed-8027287;  
 RA Driuliche B.D., Brock A., Dembait M., Yennish L., Dyson N.;  
 RT "DNA-binding and trans-activation properties of Drosophila E2F and DP  
 RT proteins.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 91:6359-6363(1994).

RL [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94158833; PubMed-8114698;  
 RA Ohtani K., Nevins J.R.;  
 RT "Functional properties of a Drosophila homolog of the E2F1 gene.";  
 RT Mol. Cell. Biol. 14:1603-1612(1994).

CC -1- FUNCTION: TRANSCRIPTION ACTIVATOR THAT BINDS DNA COOPERATIVELY  
 WITH DP PROTEIN THROUGH THE E2 RECOGNITION SITE, TTTC/GCGC,  
 FOUND IN THE PROMOTER REGION OF A NUMBER OF GENES WHOSE PRODUCTS  
 ARE INVOLVED IN CELL CYCLE REGULATION OR IN DNA REPLICATION. THE  
 DBP/E2F COMPLEX FUNCTIONS IN THE CONTROL OF CELL-CYCLE  
 PROGRESSION FROM G1 TO S PHASE. E2F-3 BINDS SPECIFICALLY TO RB1  
 PROTEIN, IN A CELL-CYCLE DEPENDENT MANNER.  
 CC -1- SUBUNIT: COMPONENT OF THE DBP1/E2F TRANSCRIPTION FACTOR COMPLEX.  
 CC BINDS COOPERATIVELY WITH DB-1 TO E2F SITES. INTERACTS WITH  
 CC RETINOBLASTOMA PROTEIN RB1 AND RELATED PROTEINS (SUCH AS P107)  
 CC THAT INHIBIT THE E2F TRANSACTIVATION DOMAIN.  
 CC -1- SUBCELLULAR LOCATION: Nucleus.  
 CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way





Wed Aug 28 13:32:30 2002

us-09-900-147-1.rsp

---

Page 12

Wed Aug 28 13:32:29 2002

US-09-900-147-1.ral

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: August 28, 2002, 12:57:09 ; Search time 12.99 seconds  
(without alignments)

69,573 Million cell updates/sec

Title: US-09-900-147-1

Perfect score: 186  
Sequence: 1 KNIRRYDALTAMNTSEKKEKIKGLPTNSA 37

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2.6/pdata/2/1aa/3A.COMB.pep.\*  
2: /cgn2.6/pdata/2/1aa/5B.COMB.pep.\*  
3: /cgn2.6/pdata/2/1aa/6A.COMB.pep.\*  
4: /cgn2.6/pdata/2/1aa/6B.COMB.pep.\*  
5: /cgn2.6/pdata/2/1aa/6C.COMB.pep.\*  
6: /cgn2.6/pdata/2/1aa/6D.COMB.pep.\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	186	100.0	37	US-09-308-935-1	Sequence 1, Appl1
2	186	100.0	72	US-08-428-131-11	Sequence 11, Appl1
3	186	100.0	72	US-09-078-596-11	Sequence 11, Appl1
4	186	100.0	410	US-08-723-415B-10	Sequence 10, Appl1
5	186	100.0	410	US-08-723-415B-11	Sequence 11, Appl1
6	186	100.0	410	US-08-428-131-2	Sequence 2, Appl1
7	186	100.0	410	US-08-602-846-2	Sequence 2, Appl1
8	186	100.0	410	US-09-078-596-2	Sequence 2, Appl1
9	186	100.0	410	US-08-189-627A-10	Sequence 10, Appl1
10	186	100.0	410	US-08-189-627A-11	Sequence 11, Appl1
11	186	100.0	369	US-08-723-415B-4	Sequence 4, Appl1
12	186	100.0	369	US-08-189-627A-4	Sequence 4, Appl1
13	186	100.0	370	US-08-723-415B-6	Sequence 6, Appl1
14	186	100.0	370	US-08-189-627A-6	Sequence 6, Appl1
15	186	100.0	385	US-08-723-415B-8	Sequence 8, Appl1
16	186	100.0	385	US-09-189-627A-8	Sequence 8, Appl1
17	186	100.0	446	US-08-723-415B-2	Sequence 2, Appl1
18	186	100.0	446	US-09-189-627A-2	Sequence 2, Appl1
19	186	100.0	446	US-08-723-415B-4	Sequence 4, Appl1
20	186	100.0	446	US-09-189-627A-4	Sequence 4, Appl1
21	186	100.0	446	US-08-723-415B-6	Sequence 6, Appl1
22	186	100.0	446	US-09-189-627A-6	Sequence 6, Appl1
23	186	100.0	446	US-08-723-415B-8	Sequence 8, Appl1
24	186	100.0	446	US-09-189-627A-8	Sequence 8, Appl1
25	186	100.0	446	US-08-723-415B-2	Sequence 2, Appl1
26	186	100.0	446	US-09-189-627A-2	Sequence 2, Appl1
27	186	100.0	446	US-08-723-415B-4	Sequence 4, Appl1

28	76	40.9	16	4	US-09-308-935-5	Sequence 5, Appl1
29	71.5	38.4	73	2	US-08-428-131-12	Sequence 12, Appl1
30	71.5	38.4	73	4	US-09-078-596-12	Sequence 12, Appl1
31	71.5	38.4	196	2	US-08-481-814A-9	Sequence 9, Appl1
32	71.5	38.4	437	1	US-08-136-119-4	Sequence 4, Appl1
33	71.5	38.4	437	1	US-07-682-711-2	Sequence 2, Appl1
34	71.5	38.4	437	2	US-08-723-415B-13	Sequence 13, Appl1
35	71.5	38.4	437	2	US-08-481-814A-6	Sequence 6, Appl1
36	71.5	38.4	437	2	US-08-162-174-2	Sequence 2, Appl1
37	71.5	38.4	437	3	US-08-801-093-1	Sequence 1, Appl1
38	71.5	38.4	437	4	US-09-189-627A-13	Sequence 13, Appl1
39	71.5	38.4	476	2	US-08-139-937-14	Sequence 14, Appl1
40	71.5	38.4	476	5	PCT-0593-11310-14	Sequence 14, Appl1
41	69.5	37.4	437	1	US-08-136-119-2	Sequence 2, Appl1
42	69.5	37.4	437	2	US-08-481-814A-7	Sequence 7, Appl1
43	69.5	37.1	14	4	US-09-308-935-11	Sequence 11, Appl1
44	65.5	35.2	413	2	US-08-481-814A-8	Sequence 8, Appl1
45	65.5	35.2	413	3	US-08-836-582-2	Sequence 2, Appl1

#### ALIGNMENTS

*Surgeal steel*

RESULT 1  
US-09-308-935-1  
Sequence 1, Application US/09308935  
Patent No. 6268334

GENERAL INFORMATION:  
APPLICANT: Bandara, Nicholas B

TITLE OF INVENTION: Peptide antagonists of DP transcription factors

FILE REFERENCE: 620-67

CURRENT FILING DATE: 1999-05-27

EARLIER APPLICATION NUMBER: PCT/GB97/03506

EARLIER FILING DATE: 1997-12-22

EARLIER APPLICATION NUMBER: GB 9626569.7

NUMBER OF SEQ ID NOS: 18

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 1

LENGTH: 37

TYPE: PPT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-308-935-1

Query Match

Best Local Similarity 100.0%; Score 186; DB 4; Length 37;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KNIRRYDALTAMNTSEKKEKIKGLPTNSA 37

1 KNIRRYDALTAMNTSEKKEKIKGLPTNSA 37

US-08-428-131-11

Sequence 11, Application US/08428131

Patent No. 5863157

GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas Barrie  
TITLE OF INVENTION: Transcription Factor DP-1  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Nixon & Vandertoye  
STREET: 1100 No. 586375th Glebe Road, 8th floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-428-131-11

Query Match 100.0%; Score 186; DB 2; Length 72;  
Best Local Similarity 100.0%; Pred. No. 3e-21;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

07 1 KNIRRYDALVLAAMNISKKEIKIGLPTNSA 37  
DB 4 KNIRRYDALVLAAMNISKKEIKIGLPTNSA 40

*sub body*

US-09-078-596-11  
Sequence 11, Application US/09078536  
Patent No. 6150116  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas Barrie  
TITLE OF INVENTION: Transcription Factor DP-1  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderrhye  
STREET: 1100 No. 6150116th Glebe Road, 8th floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/078,596  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-078-596-11

Query Match 100.0%; Score 186; DB 4; Length 72;  
Best Local Similarity 100.0%; Pred. No. 3e-21;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

07 1 KNIRRYDALVLAAMNISKKEIKIGLPTNSA 37  
DB 4 KNIRRYDALVLAAMNISKKEIKIGLPTNSA 40

US-08-723-4158-10  
Sequence 10, Application US/087234158  
Patent No. 5859198  
GENERAL INFORMATION:  
APPLICANT: Lathangue, Nicholas B.  
TITLE OF INVENTION: Transcription Factor DP-3 AND ISOFORMS  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERRHYE P.C.  
STREET: 1100 No. 5859198th Glebe Rd. 8th floor  
CITY: Arlington  
STATE: VA  
COUNTRY: USA  
ZIP: 22201-4741  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/723,4158  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: 415  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610195.1  
FILING DATE: 15-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford, Arthur R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-220  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4100  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 410 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-723-4158-10

Query Match 100.0%; Score 186; DB 2; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2.3e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

07 1 KNIRRYDALVLAAMNISKKEIKIGLPTNSA 37  
DB 163 KNIRRYDALVLAAMNISKKEIKIGLPTNSA 199

US-08-723-4158-11  
Sequence 11, Application US/087234158

Wed Aug 28 13:32:29 2002

us-09-900-147-1.ral

Page 3

Patent No. 5859199  
GENERAL INFORMATION:  
APPLICANT: LATHAQUE, Nicholas B.  
APPLICANT: delaluna, Susan  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NO. 5859199th Globe Rd. 8th floor  
CITY: Arlington  
STATE: VA  
COUNTRY: USA  
ZIP: 22201-4741  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/723,415B  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610195.1  
FILING DATE: 15-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford, Arthur R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-220  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 410 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-723-415B-11

Query Match 100.0%; Score 186; DB 2; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2,3e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

01 1 KNIRRVYDALVLMNNISKEKIKIGLPTNSA 37  
DB 163 KNIRRVYDALVLMNNISKEKIKIGLPTNSA 196

RESULT 6  
US-08-428-131-2  
Sequence 2, Application US/08428131  
Patent No. 5863757  
GENERAL INFORMATION:  
APPLICANT: LA THANGUE, Nicholas Barrie  
TITLE OF INVENTION: Transcription Factor DP-1  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye  
STREET: 1100 NO. 5863757th Globe Road, 8th floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 410 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-428-131-2

Query Match 100.0%; Score 186; DB 2; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2,3e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

01 1 KNIRRVYDALVLMNNISKEKIKIGLPTNSA 37  
DB 163 KNIRRVYDALVLMNNISKEKIKIGLPTNSA 199

RESULT 7  
US-08-602-846-2  
Sequence 2, Application US/08602846  
Patent No. 5871901  
GENERAL INFORMATION:  
APPLICANT: LA THANGUE, Nicholas B.  
TITLE OF INVENTION: ASSAY FOR INHIBITORS OF DP-1 AND OTHER DP  
TITLE OF INVENTION: PROTEINS.  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye P.C.  
STREET: 8th floor, 1100 No. 5871901th Globe Road  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/602,846  
FILING DATE: 26-FEB-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAWFORD, ARTHUR R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 620-12  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 410 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-602-846-2

Query Match 100.0%; Score 186; DB 2; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2,3e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Wed Aug 28 13:32:29 2002

us-09-900-147-1.fai

Page 4

QY 1 KNRIRRYDALNVLAMNIISEKKEIKWIGLPTNSA 37  
DB 163 KNRIRRYDALNVLAMNIISEKKEIKWIGLPTNSA 199

US-09-078-596-2  
Sequence 2, Application US/09078596

GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas Barry  
TITLE OF INVENTION: Transcription Factor DP-1  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderype  
STREET: 1100 No. 6150116th Glebe Road, 8th floor  
CITY: Arlington  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/078,596  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ALTHUR R. CLAYFORD  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 410 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-078-596-2

Query Match 100.0%; Score 186; DB 4; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2,3e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNRIRRYDALNVLAMNIISEKKEIKWIGLPTNSA 37  
DB 163 KNRIRRYDALNVLAMNIISEKKEIKWIGLPTNSA 199

US-09-189-627A-10  
Sequence 10, Application US/09189627A

GENERAL INFORMATION:  
APPLICANT: de la Lanza, Susana  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
FILE REFERENCE: 620-54  
CURRENT APPLICATION NUMBER: US/09/189,627A  
CURRENT FILING DATE: 1998-11-10  
PRIOR APPLICATION NUMBER: 08/723,415  
PRIOR FILING DATE: 1996-09-30  
PRIOR APPLICATION NUMBER: GB 9610195  
PRIOR FILING DATE: 1996-05-15  
NUMBER OF SEQ ID NOS: 25

SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 10  
LENGTH: 410  
TYPE: PRT  
ORGANISM: human  
US-09-189-627A-10

Query Match 100.0%; Score 186; DB 4; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2,3e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNRIRRYDALNVLAMNIISEKKEIKWIGLPTNSA 37  
DB 163 KNRIRRYDALNVLAMNIISEKKEIKWIGLPTNSA 199

US-09-189-627A-11  
Sequence 11, Application US/09189627A

GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
FILE REFERENCE: 620-54  
CURRENT APPLICATION NUMBER: US/09/189,627A  
CURRENT FILING DATE: 1998-11-10  
PRIOR APPLICATION NUMBER: 08/723,415  
PRIOR FILING DATE: 1996-09-30  
PRIOR APPLICATION NUMBER: GB 9610195  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 11  
LENGTH: 410  
TYPE: PRT  
ORGANISM: mouse  
US-09-189-627A-11

Query Match 100.0%; Score 186; DB 4; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2,3e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNRIRRYDALNVLAMNIISEKKEIKWIGLPTNSA 37  
DB 163 KNRIRRYDALNVLAMNIISEKKEIKWIGLPTNSA 199

US-08-723-415B-1  
Sequence 4, Application US/08723415B

GENERAL INFORMATION:  
APPLICANT: de la Lanza, Susana  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
FILE REFERENCE: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERYPE P.C.  
STREET: 1100 No. 585199th Glebe Rd. 8th floor  
CITY: Arlington  
STATE: VA  
COUNTRY: USA  
ZIP: 22201-4741  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

Wed Aug 28 13:32:29 2002

us-09-900-147-1.ra1

Page 5

APPLICATION NUMBER: US/08/723,415B  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610195.1  
FILING DATE: 15-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford, Arthur R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-220  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4100  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 369 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-723-415B-4

Query Match 97.8%; Score 182; DB 2; Length 369;  
Best Local Similarity 97.3%; Pred. No. 8,4e-20;  
Matches 36; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 101 ENIRRYDALVLMANNISKEKEIKWIGLPTNSA 137  
RESULT 12  
US-09-189-627A-4  
Sequence 4, Application US/09189627A  
Patent No. 6159691  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
FILE REFERENCE: 620-54  
CURRENT APPLICATION NUMBER: US/09/189,627A  
PRIOR FILING DATE: 1998-11-10  
PRIOR APPLICATION NUMBER: 08/723,415  
PRIOR FILING DATE: 1996-09-30  
PRIOR APPLICATION NUMBER: GB 9610195  
PRIOR FILING DATE: 1996-05-15  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: Patentn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 369  
TYPE: PRT  
ORGANISM: mouse  
US-09-189-627A-4

Query Match 97.8%; Score 182; DB 4; Length 369;  
Best Local Similarity 97.3%; Pred. No. 8,4e-20;  
Matches 36; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHUYE P.C.  
STREET: 1100 NO. 5859199th Glebe Rd. 8th floor  
CITY: Arlington  
STATE: VA  
COUNTRY: USA  
ZIP: 22201-4741  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release 1.0, Version 01.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/723,415B  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610195.1  
FILING DATE: 15-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford, Arthur R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-220  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4100  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 370 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-723-415B-6

Query Match 97.8%; Score 182; DB 2; Length 370;  
Best Local Similarity 97.3%; Pred. No. 8,4e-20;  
Matches 36; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 102 ENIRRYDALVLMANNISKEKEIKWIGLPTNSA 138  
RESULT 14  
US-09-189-627A-6  
Sequence 6, Application US/09189627A  
Patent No. 6159691  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
FILE REFERENCE: 620-54  
CURRENT APPLICATION NUMBER: US/09/189,627A  
PRIOR FILING DATE: 1998-11-10  
PRIOR APPLICATION NUMBER: 08/723,415  
PRIOR FILING DATE: 1996-09-30  
PRIOR APPLICATION NUMBER: GB 9610195  
PRIOR FILING DATE: 1996-05-15  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: Patentn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 370  
TYPE: PRT  
ORGANISM: mouse  
US-09-189-627A-6

Query Match 97.8%; Score 182; DB 4; Length 370;  
Best Local Similarity 97.3%; Pred. No. 8,4e-20;  
Matches 36; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Wed Aug 28 13:32:29 2002

us-09-900-147-1.rai

Page 6

07 1 KNIRRYDANVAMNIIISKEKINIGLPTNSA 37  
DB 102 ENIRRYDANVAMNIIISKEKINIGLPTNSA 138

RESULT 13  
US-08-723-4158-8

Sequence 8, Application US/087234158  
Patent No. 5859199

GENERAL INFORMATION:

APPLICANT: LATHROUE, Nicholas B.

APPLICANT: delaluna, Susana

TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOPORNS

TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHAYE P.C.

STREET: 1100 NO. 5859199th Globe Rd. 8th floor

CITY: Arlington

STATE: VA

COUNTRY: USA

ZIP: 22201-4741

COMPUTER READABLE FORM:

MEDIA TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/723,415B

FILING DATE: 30-SEP-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9610195.1

FILING DATE: 15-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Crawford, Arthur R.

REGISTRATION NUMBER: 25,327

REFERENCE/DOCKET NUMBER: 117-220

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-816-4000

TELEFAX: 703-816-4100

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 385 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-723-4158-8

102 e

Query Match 97.8% Score 182, DB 2, Length 385;  
Best Local Similarity 97.3% Pred. No. 8, 8e-20;  
Matches 36; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
07 1 KNIRRYDANVAMNIIISKEKINIGLPTNSA 37  
DB 117 ENIRRYDANVAMNIIISKEKINIGLPTNSA 153

Search completed: August 28, 2002, 12:59:33  
Job time: 144 sec



Wed Aug 28 13:32:29 2002

us-09-900-147-1.ral

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 12:57:09 ; Search time 12.99 Seconds  
(without alignments)  
69.573 Million cell updates/sec

Title: US-09-900-147-1

Perfect score: 186  
Sequence: 1 KNIRRYTDLVLMANNISKEKEIKWIGLPTNSA 37Scoring table: BLOSUM62  
Gapop 10.0, Capext 0.5

.ched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summariesDatabase: Issued\_Patents\_AA:  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:  
5: /cgn2\_6/ptodata/2/1aa/PCVUS.COMB.pep:  
6: /cgn2\_6/ptodata/2/1aa/Backfillal.pep:Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	186	100.0	37	4	US-09-308-935-1
2	186	100.0	72	2	US-08-428-131-11
3	186	100.0	72	4	US-09-078-596-11
4	186	100.0	410	2	US-08-723-415B-13
5	186	100.0	410	2	US-08-723-415B-11
6	186	100.0	410	2	US-08-428-131-2
7	186	100.0	410	2	US-08-602-846-2
8	186	100.0	410	4	US-09-078-596-2
9	186	100.0	410	4	US-09-189-627A-10
10	186	100.0	410	4	US-09-189-627A-11
11	186	97.8	369	2	US-08-723-415B-4
12	182	97.8	370	2	US-08-723-415B-6
13	182	97.8	370	2	US-09-189-627A-6
14	182	97.8	385	2	US-08-723-415B-8
15	182	97.8	385	4	US-09-189-627A-8
16	182	97.8	446	2	US-08-723-415B-2
17	182	97.8	446	4	US-09-189-627A-2
18	182	97.8	446	4	US-09-308-935-6
19	182	97.8	446	4	US-09-308-935-4
20	182	97.8	446	4	US-09-308-935-3
21	182	97.8	446	4	US-09-308-935-15
22	182	97.8	446	4	US-09-308-935-13
23	182	97.8	446	4	US-08-428-131-13
24	182	97.8	446	4	US-09-078-596-13
25	182	97.8	446	4	US-09-308-935-10
26	182	97.8	446	4	US-09-308-935-17
27	182	97.8	446	4	US-09-308-935-16

28	76	40.9	16	4	US-09-308-935-5	Sequence 5, Appl
29	71.5	38.4	73	2	US-08-428-131-12	Sequence 12, Appl
30	71.5	38.4	73	4	US-09-078-596-12	Sequence 12, Appl
31	71.5	38.4	196	2	US-08-481-814A-9	Sequence 9, Appl
32	71.5	38.4	437	1	US-08-136-119-4	Sequence 4, Appl
33	71.5	38.4	437	1	US-07-882-711-2	Sequence 2, Appl
34	71.5	38.4	437	2	US-08-723-415B-13	Sequence 13, Appl
35	71.5	38.4	437	2	US-08-481-814A-6	Sequence 6, Appl
36	71.5	38.4	437	2	US-08-462-174-2	Sequence 2, Appl
37	71.5	38.4	437	2	US-08-601-092-1	Sequence 1, Appl
38	71.5	38.4	437	4	US-08-189-627A-13	Sequence 13, Appl
39	71.5	38.4	476	2	US-08-139-937-14	Sequence 14, Appl
40	71.5	38.4	476	5	PCT-US93-11310-14	Sequence 14, Appl
41	69.5	37.4	437	1	US-08-136-119-2	Sequence 2, Appl
42	69.5	37.4	437	2	US-08-481-814A-7	Sequence 7, Appl
43	69	37.1	14	4	US-09-308-935-11	Sequence 11, Appl
44	65.5	35.2	413	2	US-08-481-814A-8	Sequence 8, Appl
45	65.5	35.2	413	2	US-08-836-582-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1 *Summary Start*  
US-09-308-935-1  
Sequence 1, Application US/09308935

Patent No. 6268334

GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas B

TITLE OF INVENTION: Peptide antagonists of DP transcription factors

FILE REFERENCE: C30-67

CURRENT FILING DATE: 1999-05-27

EARLIER FILING DATE: 1997-12-23

EARLIER APPLICATION NUMBER: G2 9626589.7

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 1

LENGTH: 37

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-308-935-1

Query Match 100.0%; Score 186; DB 4; Length 37;

Best Local Similarity 100.0%; Pred. No. 1.3e-21;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 KNIRRYTDLVLMANNISKEKEIKWIGLPTNSA 37

RESULT 2 *Claims DNA*  
US-08-428-131-11

Sequence 11, Application US/08428131

Patent No. 5863757

GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas Barrie

TITLE OF INVENTION: Transcription Factor DP-1

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:  
ADDRESS: Niton & Vandaele  
STREET: 1100 No. 5863757th Glebe Road, 8th floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714

Wed Aug 28 13:32:29 2002

us-09-900-147-1.fai

Page 2

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (PRO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4100  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72 amino acids  
TYPE: amino acid  
STRANDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
08-428-131-11

Query Match 100.0% Score 186; DB 2; Length 72;  
Best Local Similarity 100.0%; Pred. No. 3e-21;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 KNIRRYDALNVLMNNISKEKEINWGLPTNSA 37  
4 KNIRRYDALNVLMNNISKEKEINWGLPTNSA 40

RESULT 3 *but it had*  
US-09-078-596-11

Sequence 11, Application US/09078536

Patent No. 6150116

GENERAL INFORMATION:

APPLICANT: La Thangue, Nicholas Barrie

TITLE OF INVENTION: Transcription Factor DP-1

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSER: Nixon & Vandervee

STREET: 1100 No. 6150116th Glade Road, 8th floor

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (PRO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/078,596

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Arthur R. Crawford

REGISTRATION NUMBER: 25,327

REFERENCE/DOCKET NUMBER: 117-181

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 72 amino acids

TYPE: amino acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

STRANDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-078-596-11

Query Match 100.0% Score 186; DB 4; Length 72;  
Best Local Similarity 100.0%; Pred. No. 3e-21;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 KNIRRYDALNVLMNNISKEKEINWGLPTNSA 37  
4 KNIRRYDALNVLMNNISKEKEINWGLPTNSA 40

RESULT 4  
US-08-723-4158-10

Sequence 10, Application US/087234158

Patent No. 5859199

GENERAL INFORMATION:

APPLICANT: La Thangue, Nicholas B.

TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSER: NIXON & VANDERVEE P.C.

STREET: 1100 No. 5859199th Glade Rd, 8th floor

CITY: Arlington

STATE: VA

COUNTRY: USA

ZIP: 22201-4741

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/723,4158

FILING DATE: 30-SEP-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9610195.1

FILING DATE: 15-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Crawford, Arthur R.

REGISTRATION NUMBER: 25,327

REFERENCE/DOCKET NUMBER: 117-220

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-816-4000

TELEFAX: 703-816-4100

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 410 amino acids

TYPE: amino acid

STRANDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

Query Match 100.0% Score 186; DB 2; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2.3e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 KNIRRYDALNVLMNNISKEKEINWGLPTNSA 37  
163 KNIRRYDALNVLMNNISKEKEINWGLPTNSA 199

RESULT 5

US-08-723-4158-11

Sequence 11, Application US/087234158

Patent No. 5839139  
GENERAL INFORMATION:  
APPLICANT: LaThangue, Nicholas B.  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 No. 5859199th Glebe Rd. 8th floor  
CITY: Arlington  
STATE: VA USA  
COUNTRY: USA  
ZIP: 22201-4741  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/723,415B  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610195.1  
FILING DATE: 15-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford, Arthur R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-220  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 410 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-723-415B-11

Query Match 100.0% Score 186; DB 2; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2,3e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KNIRRYDALVLAAMNIISEKKEIKNIQLPNTSA 37  
163 KNIRRYDALVLAAMNIISEKKEIKNIQLPNTSA 159

RESULT 6  
US-08-428-131-2  
Sequence 2, Application US/08428131  
Patent No. 5863257  
GENERAL INFORMATION:  
APPLICANT: LaThangue, Nicholas BATTLE  
TITLE OF INVENTION: Transcription Factor DP-1  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE  
STREET: 1100 No. 5863757th Glebe Road, 8th floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.23 (EPO)  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
CLASSIFICATION: 518  
ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4100  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 410 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-428-131-2

Query Match 100.0% Score 186; DB 2; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2,3e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KNIRRYDALVLAAMNIISEKKEIKNIQLPNTSA 37  
163 KNIRRYDALVLAAMNIISEKKEIKNIQLPNTSA 159

RESULT 7  
US-08-602-846-2  
Sequence 2, Application US/08602846  
Patent No. 5871501  
GENERAL INFORMATION:  
APPLICANT: LaThangue, Nicholas B  
TITLE OF INVENTION: ASSAY FOR INHIBITORS OF DP-1 AND OTHER DP  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE PC  
STREET: 8th floor, 1100 No. 5871901th Glebe Road  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/602,846  
FILING DATE: 26-FEB-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAWFORD, ARTHUR R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 620-12  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4100  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 410 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-602-846-2

Query Match 100.0% Score 186; DB 2; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2,3e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Wed Aug 28 13:32:29 2002

us-09-900-147-1.ral

Page 4

QY 1 KNIRRYDALVAMNISKKEIKNIGLPTNSA 37  
DB 163 KNIRRYDALVAMNISKKEIKNIGLPTNSA 199

RESULT 8

US-09-078-596-2  
Sequence 2, Application US/09078596

PATENT NO. 6150116  
GENERAL INFORMATION:  
APPLICANT: LA THANGUE, NICHOLAS BARRIE  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-1  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHVE  
STREET: 1100 NO. 6150116TH GLEBE ROAD, 8th FLOOR  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/078,596  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ARTHUR R. CRAWFORD  
REGISTRATION NUMBER: 25,317  
REFERENCE/DOCKET NUMBER: 117-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 410 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-078-596-2

Query Match 100.0%; Score 186; DB 4; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2,3e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNIRRYDALVAMNISKKEIKNIGLPTNSA 37  
DB 163 KNIRRYDALVAMNISKKEIKNIGLPTNSA 199

RESULT 9

US-09-189-627A-10  
Sequence 10, Application US/09189627A  
Patent No. 6159691

GENERAL INFORMATION:  
APPLICANT: LA THANGUE, NICHOLAS  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
FILE REFERENCE: 620-54  
CURRENT APPLICATION NUMBER: US/09/189,627A  
PRIOR FILING DATE: 1998-11-10  
PRIOR APPLICATION NUMBER: 08/723,415  
PRIOR FILING DATE: 1996-09-30  
PRIOR APPLICATION NUMBER: GB 9610195  
PRIOR FILING DATE: 1996-05-15  
NUMBER OF SEQ ID NOS: 25

SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 10  
LENGTH: 410  
TYPE: PRT  
ORGANISM: human  
US-09-189-627A-10

Query Match 100.0%; Score 186; DB 4; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2,3e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNIRRYDALVAMNISKKEIKNIGLPTNSA 37  
DB 163 KNIRRYDALVAMNISKKEIKNIGLPTNSA 199

RESULT 10

US-09-189-627A-11  
Sequence 11, Application US/09189627A  
Patent No. 6159691

GENERAL INFORMATION:  
APPLICANT: LA THANGUE, NICHOLAS  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
FILE REFERENCE: 620-54  
CURRENT APPLICATION NUMBER: US/09/189,627A  
CURRENT FILING DATE: 1998-11-10  
PRIOR APPLICATION NUMBER: 08/723,415  
PRIOR FILING DATE: 1996-09-30  
PRIOR APPLICATION NUMBER: GB 9610195  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 11  
LENGTH: 410  
TYPE: PRT  
ORGANISM: mouse  
US-09-189-627A-11

Query Match 100.0%; Score 186; DB 4; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2,3e-20;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNIRRYDALVAMNISKKEIKNIGLPTNSA 37  
DB 163 KNIRRYDALVAMNISKKEIKNIGLPTNSA 199

RESULT 11

US-08-723-415B  
Sequence 4, Application US/08723415B  
Patent No. 5859199

GENERAL INFORMATION:  
APPLICANT: LA THANGUE, NICHOLAS B.  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHVE P.C.  
STREET: 1100 NO. 5859199TH GLEBE RD. 8th FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22201-4741  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

Wed Aug 28 13:32:29 2002

us-09-900-147-1.rat

Page 5

APPLICATION NUMBER: US/08/723,415B  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610195.1  
FILING DATE: 15-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford, Arthur R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-220  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 369 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-723-415B-4

Query Match 97.8% Score 182; DB 2; Length 369;  
Best Local Similarity 97.3% Pred. No. 8.4e-20;  
Matches 36; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 101 ENIRRYDALVLMNNISKEKEIKWIGLPTNSA 37  
:|||||

RESULT 12  
US-09-189-627A-4  
Sequence 4, Application US/09189627A  
Patent No. 6139691  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas  
APPLICANT: de la Luna, Susana  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
FILE REFERENCE: 620-54  
CURRENT APPLICATION NUMBER: US/09/189,627A  
PRIOR FILING DATE: 1996-11-10  
PRIOR APPLICATION NUMBER: 08/723,415  
PRIOR FILING DATE: 1996-09-30  
PRIOR APPLICATION NUMBER: GB 9610195  
PRIOR FILING DATE: 1996-05-15  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 369  
TYPE: PRT  
ORGANISM: mouse  
US-09-189-627A-4

Query Match 97.8% Score 182; DB 4; Length 369;  
Best Local Similarity 97.3% Pred. No. 8.4e-20;  
Matches 36; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 101 ENIRRYDALVLMNNISKEKEIKWIGLPTNSA 37  
:|||||

RESULT 13  
US-08-723-415B-6  
Sequence 6, Application US/08723415B  
Patent No. 5859199  
GENERAL INFORMATION:  
APPLICANT: LaThangue, Nicholas B.  
APPLICANT: delaluna, Susana  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS  
THEREOF

NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERBYE P.C.  
STREET: 1100 NO. 5859199th globe rd. 8th floor  
CITY: Arlington  
STATE: VA  
COUNTRY: USA  
ZIP: 22201-4741  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release 1.0, Version 91.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/723,415B  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610195.1  
FILING DATE: 15-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford, Arthur R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-220  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 370 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-723-415B-6

Query Match 97.8% Score 182; DB 2; Length 370;  
Best Local Similarity 97.3% Pred. No. 8.4e-20;  
Matches 36; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 102 ENIRRYDALVLMNNISKEKEIKWIGLPTNSA 138  
:|||||

RESULT 14  
US-09-189-627A-6  
Sequence 6, Application US/09189627A  
Patent No. 6139691  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas  
APPLICANT: de la Luna, Susana  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
FILE REFERENCE: 620-54  
CURRENT APPLICATION NUMBER: US/09/189,627A  
PRIOR FILING DATE: 1996-11-10  
PRIOR APPLICATION NUMBER: 08/723,415  
PRIOR FILING DATE: 1996-09-30  
PRIOR APPLICATION NUMBER: GB 9610195  
PRIOR FILING DATE: 1996-05-15  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 370  
TYPE: PRT  
ORGANISM: mouse  
US-09-189-627A-6

Query Match 97.8% Score 182; DB 4; Length 370;  
Best Local Similarity 97.3% Pred. No. 8.4e-20;  
Matches 36; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Wed Aug 28 13:32:29 2002

us-09-900-147-1.rpt

Page 6

07 1 KNIRRYDALNVLAMNIIISKEKEIKWIGLPTNSA 37  
DB 102 ENIRRYDALNVLAMNIIISKEKEIKWIGLPTNSA 138  
RESULT 15  
US-08-723-4158-8  
Sequence 8, Application US/087234158  
Patent No. 5859199  
GENERAL INFORMATION:  
APPLICANT: Lazarene, Nicholas B.  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHAYE P.C.  
STREET: 1100 No. 5859199th Glebe Rd. 8th floor  
CITY: Arlington  
STATE: VA  
COUNTRY: USA  
ZIP: 22201-4741  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/723-4158  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: A35  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610195.1  
FILING DATE: 13-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford, Arthur R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-220  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 385 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-723-4158-8

102 e

Query Match 97.8% Score 182; DB 2; Length 385;  
eat Local Similarity 97.3%; Pred. No. 8.8e-20;  
atches 36; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
07 1 KNIRRYDALNVLAMNIIISKEKEIKWIGLPTNSA 37  
DB 117 ENIRRYDALNVLAMNIIISKEKEIKWIGLPTNSA 153

Search completed: August 28, 2002, 12:59:33  
Job time: 144 sec